GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 10:41:02 ; Search time 131.44 Seconds
(without alignments)
6914.267 Million cell updates/sec

Title: US-09-909-474D-1
Perfect score: 2059
Sequence: 1 gctctgagccccggcggcgc.....aaaaaaaaaaaaaaaaaaaa 2059

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1004.4	48.8	1085	4	US-09-509-902A-3	Sequence 3, Appli
c 2	505.2	24.5	2559	4	US-09-220-132-135	Sequence 135, App
c 3	505.2	24.5	2562	1	US-08-146-421-4	Sequence 4, Appli
4	85	4.1	1302	1	US-08-913-050A-2	Sequence 2, Appli
5	85	4.1	1302	4	US-09-016-434-1146	Sequence 1146, Ap
6	84.6	4.1	1545	4	US-09-559-023-1	Sequence 1, Appli
7	82.8	4.0	1674	4	US-09-996-243-300	Sequence 300, App
8	82.4	4.0	2136	4	US-09-996-243-302	Sequence 302, App
9	81.6	4.0	1466	2	US-08-749-902-2	Sequence 2, Appli
10	79.8	3.9	1048	4	US-09-489-847-38	Sequence 38, Appl
11	79.6	3.9	485	4	US-09-996-243-245	Sequence 245, App
12	79.6	3.9	1066	1	US-08-157-101A-4	Sequence 4, Appli
13	79.6	3.9	1927	4	US-09-336-536-66	Sequence 66, Appl
14	79.2	3.8	1091	4	US-09-328-965-1	Sequence 1, Appli
15	79.2	3.8	1882	3	US-09-370-253-1	Sequence 1, Appli
16	78.8	3.8	1024	4	US-09-328-475C-50	Sequence 50, Appl
17	78.8	3.8	2674	4	US-09-817-180-1	Sequence 1, Appli
18	78.4	3.8	2246	4	US-09-363-708-3	Sequence 3, Appli
19	78.4	3.8	2246	4	US-09-083-587-3	Sequence 3, Appli
20	78.2	3.8	1342	4	US-09-489-847-89	Sequence 89, Appl
21	78.2	3.8	2434	4	US-09-489-847-67	Sequence 67, Appl
22	78	3.8	1582	3	US-08-545-196B-10	Sequence 10, Appl
23	78	3.8	1582	3	US-08-545-196B-12	Sequence 12, Appl
24	77.8	3.8	903	5	PCT-US95-06406A-21	Sequence 21, Appl
25	77.8	3.8	1872	4	US-09-801-052-1	Sequence 1, Appli
26	77.8	3.8	2323	4	US-09-149-476-24	Sequence 24, Appl
27	77.8	3.8	3715	4	US-09-234-245-1	Sequence 1, Appli
28	77.4	3.8	140	1	US-08-628-417-5	Sequence 5, Appli
29	77.4	3.8	240	1	US-08-628-417-6	Sequence 6, Appli
30	77.4	3.8	2184	3	US-08-955-918C-1	Sequence 1, Appli
31	77.4	3.8	2184	4	US-08-697-766A-1	Sequence 1, Appli
32	77.4	3.8	2186	4	US-09-360-545-66	Sequence 66, Appl
33	77.4	3.8	3060	4	US-09-996-243-504	Sequence 504, App
34	77.2	3.7	1249	4	US-09-461-325-128	Sequence 128, App
c 35	77.2	3.7	1260	4	US-09-461-325-93	Sequence 93, Appl
36	77.2	3.7	2202	4	US-09-465-558-59	Sequence 59, Appl
37	77.2	3.7	2447	2	US-09-014-969-14	Sequence 14, Appl
38	77	3.7	1198	3	US-09-248-335-27	Sequence 27, Appl
39	77	3.7	1414	4	US-09-501-115-5	Sequence 5, Appli
40	76.8	3.7	1411	3	US-08-964-127-5	Sequence 5, Appli
41	76.8	3.7	1411	4	US-09-496-692-5	Sequence 5, Appli
42	76.8	3.7	1411	4	US-10-000-273-5	Sequence 5, Appli
43	76.6	3.7	578	4	US-09-602-877A-95	Sequence 95, Appl
44	76.6	3.7	949	4	US-09-489-847-35	Sequence 35, Appl
45	76.6	3.7	1683	3	US-09-347-803-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-509-902A-3
; Sequence 3, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1085
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-3

Query Match 48.8%; Score 1004.4; DB 4; Length 1085;
Best Local Similarity 99.9%; Pred. No. 4.8e-207;
Matches 1005; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GCTCTGAGCCCCGGCGGCGCCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 139

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 199

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 259

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 319

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 379

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 439

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
```

```
Db        440 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 499
Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
Db        500 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 559
Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
Db        560 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 619
Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
Db        620 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 679
Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
Db        680 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 739
Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
Db        740 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 799
Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
Db        800 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 859
Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
Db        860 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 919
Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
Db        920 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 979
Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
Db        980 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1039
Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1006
Db       1040 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1085
```

RESULT 2
US-09-220-132-135/c
; Sequence 135, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCE
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 135
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-135

Query Match 24.5%; Score 505.2; DB 4; Length 2559;
Best Local Similarity 94.6%; Pred. No. 1.4e-99;
Matches 566; Conservative 0; Mismatches 28; Indels 4; Gaps 4;

```
Qy        604 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 663
Db       1634 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 1575
Qy        664 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 723
Db       1574 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT-GGACAAGCACGCGTCGCCAGCCTAC 1516
Qy        724 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 783
Db       1515 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 1456
Qy        784 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 843
Db       1455 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 1396
Qy        844 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 903
Db       1395 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGCGCTCTCG 1336
Qy        904 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 963
Db       1335 CCCTCGCCGGGCTGTCTGGTTCCGTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 1276
Qy        964 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 1023
Db       1275 G-CACAGCGATCCTCCTGCACCCCTGGCTGCGACA-GACCCGATGCCCTTAGCCCCAACC 1218
Qy       1024 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1083
Db       1217 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA-CC 1159
Qy       1084 AGGGAAGAGGAGGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1143
Db       1158 AGGGAAGAGGAGGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1099
Qy       1144 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGA 1201
Db       1098 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGA 1041
```

RESULT 3
US-08-146-421-4/c
; Sequence 4, Application US/08146421
; Patent No. 5541499
; GENERAL INFORMATION:
; APPLICANT: BREWER, GARY
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A POLYPEPTIDE WITH
; TITLE OF INVENTION: ANTI-TUMOR PROPERTIES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DILWORTH & BARRESE
; STREET: 4350 LA JOLLA VILLAGE DRIVE, SUITE 300
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,421
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PEPPER PH.D., FREDERICK W.
; REGISTRATION NUMBER: 31,286
; REFERENCE/DOCKET NUMBER: 489-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4410
; TELEFAX: 619-453-2839
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 14:49:15 ; Search time 366.799 Seconds
(without alignments)
10320.639 Million cell updates/sec

Title: US-09-909-474D-1_COPY_49_1122
Perfect score: 1074
Sequence: 1 atgcgagccacccctctggc.....gagaagtggttctgtatggc 1074

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1070.8	99.7	1074	9	US-09-799-875-9	Sequence 9, Appli
2	1070.8	99.7	2389	9	US-09-799-875-7	Sequence 7, Appli
3	1069.2	99.6	2092	14	US-10-098-841-271	Sequence 271, App
4	956.4	89.1	1085	15	US-10-024-828-3	Sequence 3, Appli
5	879.2	81.9	2048	9	US-09-925-301-260	Sequence 260, App
6	282.2	26.3	396	11	US-09-803-719-519	Sequence 519, App
7	263.2	24.5	1909	12	US-10-291-172-116	Sequence 116, App
8	263.2	24.5	1909	15	US-10-119-926-73	Sequence 73, Appl
9	253	23.6	4336	15	US-10-084-817-300	Sequence 300, App
10	251.4	23.4	3280	14	US-10-044-090-269	Sequence 269, App
11	247	23.0	3324	13	US-10-240-965-168	Sequence 168, App
12	232.4	21.6	818	15	US-10-228-263-1	Sequence 1, Appli
13	222.2	20.7	621	15	US-10-228-263-3	Sequence 3, Appli
14	220.8	20.6	562	9	US-09-925-301-525	Sequence 525, App
15	207.6	19.3	1943	12	US-10-291-172-492	Sequence 492, App
c 16	165	15.4	498	10	US-09-917-800A-991	Sequence 991, App
c 17	158	14.7	562	9	US-09-864-761-12987	Sequence 12987, A
c 18	156	14.5	417	9	US-09-864-761-29550	Sequence 29550, A
19	138.8	12.9	4463	12	US-10-062-674-1704	Sequence 1704, Ap
20	80	7.4	625	15	US-10-084-817-301	Sequence 301, App
c 21	78.6	7.3	552	9	US-09-864-761-12930	Sequence 12930, A
22	78.4	7.3	1444	9	US-09-867-550-1771	Sequence 1771, Ap
c 23	76.2	7.1	293	9	US-09-864-761-29493	Sequence 29493, A
24	75.4	7.0	1795	14	US-10-006-611-1	Sequence 1, Appli
25	66.8	6.2	1432	10	US-09-836-392-9	Sequence 9, Appli
26	65.6	6.1	2169	13	US-10-204-041-15	Sequence 15, Appl
27	65.6	6.1	2169	15	US-10-108-580-1	Sequence 1, Appli
28	65.4	6.1	1884	9	US-09-963-159-3	Sequence 3, Appli
29	65.4	6.1	3353	9	US-09-963-159-1	Sequence 1, Appli
30	55.6	5.2	1376	10	US-09-764-868-237	Sequence 237, App
31	55.2	5.1	1933	13	US-10-094-749-112	Sequence 112, App
32	54.6	5.1	2902	13	US-10-355-975-4	Sequence 4, Appli
33	54	5.0	819	15	US-10-275-546-1	Sequence 1, Appli
34	54	5.0	858	13	US-10-029-386-25154	Sequence 25154, A
35	54	5.0	1149	13	US-09-863-776-1	Sequence 1, Appli
c 36	54	5.0	32192	10	US-09-764-847-1416	Sequence 1416, Ap
37	54	5.0	32192	11	US-09-764-891-7945	Sequence 7945, Ap
38	54	5.0	32192	11	US-09-764-891-8220	Sequence 8220, Ap
c 39	54	5.0	32192	11	US-09-764-891-8319	Sequence 8319, Ap
c 40	54	5.0	32192	15	US-10-092-154-1416	Sequence 1416, Ap
41	53.8	5.0	1368	10	US-09-764-868-557	Sequence 557, App
42	53.8	5.0	1368	11	US-09-764-891-1495	Sequence 1495, Ap
c 43	51.8	4.8	5365	13	US-10-027-632-257387	Sequence 257387,
c 44	51.8	4.8	5365	14	US-10-027-632-257387	Sequence 257387,
45	51.6	4.8	3120	15	US-10-112-286-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-799-875-9
; Sequence 9, Application US/09799875
; Patent No. US20020034780A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. US20020034780A1el Human Protein Kinases and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-799-875-9

Query Match 99.7%; Score 1070.8; DB 9; Length 1074;
Best Local Similarity 99.8%; Pred. No. 3.7e-272;
Matches 1072; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
```

```
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360

Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Db        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Db        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Db        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840

Db        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900

Db        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960

Db        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020

Db        961 CCCTTAGCTCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074

Db       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074


RESULT 2
US-09-799-875-7
; Sequence 7, Application US/09799875
; Patent No. US20020034780A1
; GENERAL INFORMATION:
;  APPLICANT:  Meyers, Rachel
;  APPLICANT:  Kapeller-Libermann, Rosana
;  APPLICANT:  Williamson, Mark
;  TITLE OF INVENTION:  No. US20020034780A1el Human Protein Kinases and Uses
;  TITLE OF INVENTION:  Therefor
;  FILE REFERENCE:  35800/209996
;  CURRENT APPLICATION NUMBER: US/09/799,875
;  CURRENT FILING DATE:  2001-03-06
;  PRIOR APPLICATION NUMBER:  60/182,059
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  09/659,287
;  PRIOR FILING DATE:  2000-09-12
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;  LENGTH: 2389
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (383)...(1456)
US-09-799-875-7

  Query Match             99.7%;  Score 1070.8;  DB 9;  Length 2389;
  Best Local Similarity   99.8%;  Pred. No. 3.9e-272;
  Matches 1072;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Db        383 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 442

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Db        443 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 502

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Db        503 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 562

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Db        563 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 622

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Db        623 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 682

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360

Db        683 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 742

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Db        743 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 802

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Db        803 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 862

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Db        863 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 922

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Db        923 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 982

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Db        983 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 1042

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Db       1043 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 1102

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
```

```
Db       1103 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 1162

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1163 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 1222

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1223 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1282

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1283 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1342

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1343 CCCTTAGCTCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1402

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1403 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1456


RESULT 3
US-10-098-841-271
; Sequence 271, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2
;  CURRENT APPLICATION NUMBER: US/10/098,841
;  CURRENT FILING DATE:  2002-03-13
;  PRIOR APPLICATION NUMBER: 09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 331
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 271
;   LENGTH: 2092
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (139)..(1215)
US-10-098-841-271

  Query Match             99.6%;  Score 1069.2;  DB 14;  Length 2092;
  Best Local Similarity   99.7%;  Pred. No. 1e-271;
  Matches 1071;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 198

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 258

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 318

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 378

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        379 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAGGGTGTACCCC 438

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        439 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 498

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 558

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 618

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 678

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 738

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 798

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 858

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 918

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 978

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1038

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1039 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1098

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1158

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1212


RESULT 4
US-10-024-828-3
; Sequence 3, Application US/10024828
```

```
; Publication No. US20030036051A1
; GENERAL INFORMATION:
;  APPLICANT: Virca, Duke
;  APPLICANT:  Bird, Timothy A.
;  APPLICANT:  Anderson, Dirk M.
;  APPLICANT:  Marken, John S.
;  TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase
;  TITLE OF INVENTION:  Functions
;  FILE REFERENCE: 2877-US
;  CURRENT APPLICATION NUMBER: US/10/024,828
;  CURRENT FILING DATE:  2001-12-18
;  PRIOR APPLICATION NUMBER: US/09/509,902A
;  PRIOR FILING DATE: 1999-08-03
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1085
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-024-828-3

  Query Match             89.1%;  Score 956.4;  DB 15;  Length 1085;
  Best Local Similarity   99.9%;  Pred. No. 4.7e-242;
  Matches  957;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 187

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 247

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 307

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 367

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 427

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 487

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 547

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 607

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 667

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 727

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 787

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 847

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        848 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 907

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 967

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        968 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1027

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1085


RESULT 5
US-09-925-301-260
; Sequence 260, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 260
;   LENGTH: 2048
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (66)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (67)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-260

  Query Match             81.9%;  Score 879.2;  DB 9;  Length 2048;
  Best Local Similarity   92.6%;  Pred. No. 1e-221;
  Matches  994;  Conservative    3;  Mismatches    1;  Indels   76;  Gaps    4;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 215

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 275

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||
Db        276 CCCAGACTGCCCCCCTGCCTGTTGCCCCT---------------------------- 304

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
                                                  ||||||||||||||||||||
Db        305 ------------------------------------GAGCCCGAGGAGGGCGGG 322

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| || |||||||||||||||||||||||||||||||||| |||||||||
Db        323 CGGGCCTACCAGG-CCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 380

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        381 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTAYGCGCGGCTGCCCCCGCACAAGCATGTG 440
```

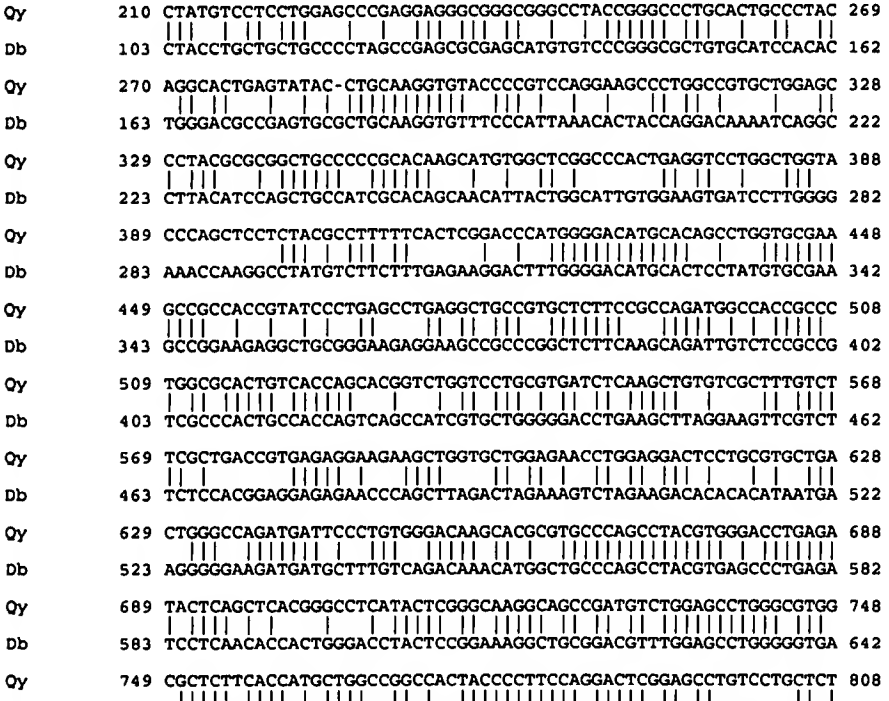
```
Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 500

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        501 CATGGGGACATGCACAGCCTGGTGCGAAG-CGCCACCGTATCCCTGAGCCTGAGGCTGCC 559

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 619

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 679

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 739

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        740 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCAMGGGCCTCATACTCGGGCAAG 799

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 859

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 919

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 979

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1039

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 CCCTTAGCYCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1099

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1153
```

```
RESULT 6
US-09-803-719-519
; Sequence 519, Application US/09803719
; Publication No. US20030044783A1
; GENERAL INFORMATION:
;  APPLICANT: Williams, Lewis T.
;  APPLICANT:  Escobedo, Jaime
;  APPLICANT:  Innis, Michael A.
;  APPLICANT:  Garcia, Pablo Dominiguez
;  APPLICANT:  Sudduth-Klinger, Julie
;  APPLICANT:  Reinhard, Christoph
;  APPLICANT:  Giese, Klaus
;  APPLICANT:  Randazzo, Filippo
;  APPLICANT:  Kennedy, Giulia C.
;  APPLICANT:  Pot, David
;  APPLICANT:  Kassam, Altaf
;  APPLICANT:  Lamson, George
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Jones, Lee William
;  APPLICANT:  Stache-Crain, Birgit
;  TITLE OF INVENTION: Human Genes and Gene Products
;  FILE REFERENCE: 1624.002
;  CURRENT APPLICATION NUMBER: US/09/803,719
;  CURRENT FILING DATE:  2001-03-09
;  PRIOR APPLICATION NUMBER: 60/188,609
;  PRIOR FILING DATE: 2000-03-09
;  NUMBER OF SEQ ID NOS: 2396
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 519
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-803-719-519

  Query Match             26.3%;  Score 282.2;  DB 11;  Length 396;
  Best Local Similarity   97.3%;  Pred. No. 1.2e-64;
  Matches  287;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGATGGAG 160

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TTGGATGACAACTTAGATACCGAGCGTTCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 220

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 280

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 340

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGT 295
              ||| |||||| ||||||||||||||||||||||||||    ||||||||||||||
Db        341 CGGTCCTACCAGGCCCTGCACTGCCCTACAGGCACTTGAGATACCTGCAAGGTGT 395
```

```
RESULT 7
US-10-291-172-116
; Sequence 116, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-045
;  CURRENT APPLICATION NUMBER: US/10/291,172
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 09/693,267
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/665,363
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: 09/616,847
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 09/596,193
;  PRIOR FILING DATE: 2000-06-17
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 116
;   LENGTH: 1909
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (135)..(941)
```

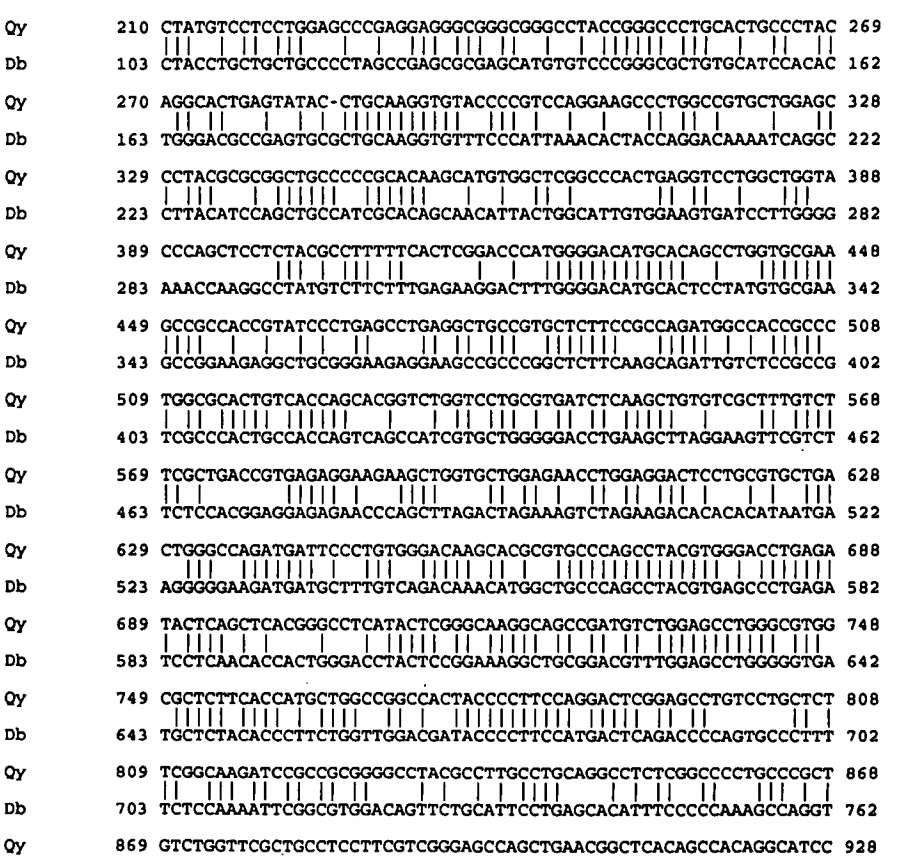
US-10-291-172-116

Query Match 24.5%; Score 263.2; DB 12; Length 1909;
Best Local Similarity 60.7%; Pred. No. 1.2e-59;
Matches 447; Conservative 0; Mismatches 288; Indels 1; Gaps 1;



RESULT 8
US-10-119-926-73
; Sequence 73, Application US/10119926
; Publication No. US20030104413A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Peiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104413A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 789CIP2BCON
; CURRENT APPLICATION NUMBER: US/10/119,926
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 73
; LENGTH: 1909
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(944)
US-10-119-926-73

Query Match 24.5%; Score 263.2; DB 15; Length 1909;
Best Local Similarity 60.7%; Pred. No. 1.2e-59;
Matches 447; Conservative 0; Mismatches 288; Indels 1; Gaps 1;



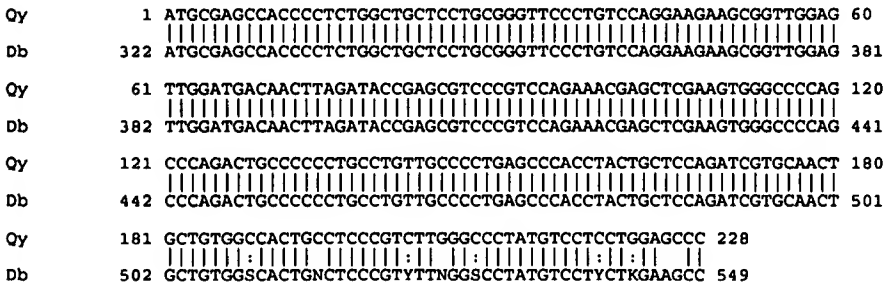
```
Db        763 GCCTCATTCGCAGCCTCTTGAGACGGGAGCCCTCCGAGAGACTCACTGCCCCCGAGATCC 822

Qy        929 TCCTGCACCCCTGGCT 944
              | |||||||||||| |
Db        823 TACTGCACCCCTGGTT 838


RESULT 9
US-10-084-817-300
; Sequence 300, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
;  APPLICANT: Susan Stuart
;  APPLICANT:  Jed G. Nuchtern
;  APPLICANT:  Sharon E. Plon
;  APPLICANT:  Jason M. Shohet
;  TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
;  FILE REFERENCE: PA-0046 US
;  CURRENT APPLICATION NUMBER: US/10/084,817
;  CURRENT FILING DATE:  2002-02-25
;  PRIOR APPLICATION NUMBER: 60/270,784
;  PRIOR FILING DATE: 2001-02-23
;  NUMBER OF SEQ ID NOS: 365
;  SOFTWARE: PERL Program
; SEQ ID NO 300
;   LENGTH: 4336
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030119009A1 234427.4
US-10-084-817-300

  Query Match           23.6%;  Score 253;  DB 15;  Length 4336;
  Best Local Similarity   55.7%;  Pred. No. 6e-57;
  Matches  484;  Conservative    0;  Mismatches  385;  Indels    0;  Gaps    0;

Qy        147 CCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGG 206
              |||  || | || |   | || ||  | | |       |      |  ||  || | ||
Db       1459 CCTCGGCTCCCCGAGCCCGCCCGAGACTCCGAACTTGTCGCATTGCGTTTCTTGTATCGG 1518

Qy        207 GCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCC 266
              |   ||  |  |  |||| ||   |||||| |  |  |  |  || ||| ||||    |
Db       1519 GAAATACTTATTGTTGGAACCTCTGGAGGGAGACCACGTTTTTCGTGCCGTGCATCTGCA 1578

Qy        267 TACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGA 326
              |  ||    |||        ||||||||||        ||        || ||    ||||
Db       1579 CAGCGGAGAGGAGCTGGTGTGCAAGGTGTTTGATATCAGCTGCTACCAGGAATCCCTGGC 1638

Qy        327 GCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGG 386
              || | |    | ||| |  | || |   |  |     |  ||||||  |  |    ||
Db       1639 ACCGTGCTTTTGCCTGTCTGCTCATAGTAACATCAACCAAATCACTGAAATTATCCTGGG 1698

Qy        387 TACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCG 446
              |          ||| |  || ||    || | | ||||||||||||||    | | || ||
Db       1699 TGAGACCAAAGCCTATGTGTTCTTTGAGCGAAGCTATGGGGACATGCATTCCTTCGTCCG 1758

Qy        447 AAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGC 506
              | | || |      |   |||  ||||| |||   || |||  |||||| ||| | ||
Db       1759 CACCTGCAAGAAGCTGAGAGAGGAGGAGGCAGCCAGACTGTTCTACCAGATTGCCTCGGC 1818

Qy        507 CCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGT 566
              ||||  ||||| ||  |    || ||||| |||||  || |||||||| |    ||  |
Db       1819 AGTGGCCCACTGCCATGACGGGGGGCTGGTGCTGCGGGACCTCAAGCTGCGGAAATTCAT 1878

Qy        567 CTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCT 626
              |||    |||  |||||||    | |   |||||| | |||||| ||| ||| | | ||
Db       1879 CTTTAAGGACGAAGAGAGGACTCGGGTCAAGCTGGAAAGCCTGGAAGACGCCTACATTCT 1938

Qy        627 GACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGA 686
              |  ||    |||||||||||  | ||||||||  | |||||  || |||||   | || ||
Db       1939 GCGGGGAGATGATGATTCCCTCTCCGACAAGCATGGCTGCCCGGCTTACGTAAGCCCAGA 1998

Qy        687 GATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGT 746
              |||  | | | |  | | |   |||||||||||| |||||||| || ||||||||||| ||
Db       1999 GATCTTGAACACCAGTGGCAGCTACTCGGGCAAAGCAGCCGACGTGTGGAGCCTGGGGGT 2058

Qy        747 GGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCT 806
              |  ||| | |||||||| |||  || |  |||||  ||||| |||   || ||   |  ||
Db       2059 GATGCTGTACACCATGTTGGTGGGGCGGTACCCTTTCCATGACATTGAACCCAGCTCCCT 2118

Qy        807 CTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCG 866
              |||| |||||||||| || ||     | |  | | || |     || ||| ||   |||
Db       2119 CTTCAGCAAGATCCGGCGTGGCCAGTTCAACATTCCAGAGACTCTGTCGCCCAAGGCCAA 2178

Qy        867 CTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCAT 926
              || ||  | ||  || | || ||||||||||  | || |||||  ||  |   |  ||
Db       2179 GTGCCTCATCCGAAGCATTCTGCGTCGGGAGCCCTCAGAGCGGCTGACCTCGCAGGAAAT 2238

Qy        927 CCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCT 986
              ||    || || ||| |        ||     |    ||       || |          |
Db       2239 TCTGGACCATCCTTGGTTTTCTACAGATTTTAGCGTCTCGAATTCAGCATATGGTGCTAA 2298

Qy        987 CTGGGAGGCTGCCCAGGTGGTCCCTGATG 1015
              | | ||| |||| |||| || || |
Db       2299 GGAAGTGTCTGACCAGCTGGTGCCGGACG 2327


RESULT 10
US-10-044-090-269
; Sequence 269, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 269
;   LENGTH: 3280
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1820904CB1
US-10-044-090-269

  Query Match           23.4%;  Score 251.4;  DB 14;  Length 3280;
  Best Local Similarity   55.6%;  Pred. No. 1.6e-56;
  Matches  483;  Conservative    0;  Mismatches  386;  Indels    0;  Gaps    0;

Qy        147 CCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGG 206
              |||  || | || |   | || ||  | | |       |      |  ||  || | ||
Db        403 CCTCGGCTCCCCGAGCCCGCCCGAGACTCCGAACTTGTCGCATTGCGTTTCTTGTATCGG 462

Qy        207 GCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCC 266
              |   ||  |  |  |||| ||   |||||| |  |  |  |  || ||| ||||    |
Db        463 GAAATACTTATTGTTGGAACCTCTGGAGGGAGACCACGTTTTTCGTGCCGTGCATCTGCA 522

Qy        267 TACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGA 326
              |  ||    |||        ||||||||||        ||        || ||    ||||
Db        523 CAGCGGAGAGGAGCTGGTGTGCAAGGTGTTTGATATCAGCTGCTACCAGGAATCCCTGGC 582

Qy        327 GCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGG 386
              || | |    | ||| |  | || |   |  |     |  ||||||  |  |    ||
Db        583 ACCGTGCTTTTGCCTGTCTGCTCATAGTAACATCAACCAAATCACTGAAATTATCCTGGG 642

Qy        387 TACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCG 446
              |          ||| |  || ||    || | | ||||||||||||||    | | || ||
Db        643 TGAGACCAAAGCCTATGTGTTCTTTGAGCGAACGTATGGGGACATGCATTCCTTCGTCCG 702

Qy        447 AAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGC 506
              | | || |      |   |||  ||||| |||   || |||  |||||| ||| | ||
```

```
Db        703 CACCTGCAAGAAGCTGAGAGAGGAGGAGGCAGCCAGACTGTTCTACCAGATTGCCTCGGC 762

Qy        507 CCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGT 566

Db        763 AGTGGCCCACTGCCATGACGGGGGGCTGGTGCTGCGGGACCTCAAGCTGCGGAAATTCAT 822

Qy        567 CTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCT 626

Db        823 CTTTAAGGACGAAGAGAGGACTCGGGTCAAGCTGGAAAGCCTGGAAGACGCCTACATTCT 882

Qy        627 GACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGA 686

Db        883 GCGGGGAGATGATGATTCCCTCTCCGACAAGCATGGCTGCCCGGCTTACGTAAGCCCAGA 942

Qy        687 GATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGT 746

Db        943 GATCTTGAACACCAGTGGCAGCTACTCGGGCAAAGCAGCCGACGTGTGGAGCCTGGGGGT 1002

Qy        747 GGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCT 806

Db       1003 GATGCTGTACACCATGTTGGTGGGGCGGTACCCTTTCCATGACATTGAACCCAGCTCCCT 1062

Qy        807 CTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCG 866

Db       1063 CTTCAGCAAGACCCGGCGTGGCCAGTTCAACATTCCAGAGACTCTGTCGCCCAAGGCCAA 1122

Qy        867 CTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCAT 926

Db       1123 GTGCCTCATCCGAAGCATTCTGCGTCGGGAGCCCTCAGAGCGGCTGACCTCGCAGGAAAT 1182

Qy        927 CCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCT 986

Db       1183 TCTGGACCATCCTTGGTTTTCTACAGATTTTAGCGTCTCGAATTCAGCATATGGTGCTAA 1242

Qy        987 CTGGGAGGCTGCCCAGGTGGTCCCTGATG 1015

Db       1243 GGAAGTGTCTGACCAGCTGGTGCCGGACG 1271


RESULT 11
US-10-240-965-168
; Sequence 168, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  SHIFFMAN, Dov
;  APPLICANT:  SOMOGYI, Roland
;  APPLICANT:  LAWN, Richard M.
;  APPLICANT:  SEILHAMER, Jeffrey J.
;  APPLICANT:  PORTER, Gordon J.
;  APPLICANT:  MIKITA, Thomas
;  APPLICANT:  TAI, Julie
;  TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
;  FILE REFERENCE: PA-0025 PCT
;  CURRENT APPLICATION NUMBER: US/10/240,965
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/195,106
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 276
;  SOFTWARE: PERL Program
; SEQ ID NO 168
;   LENGTH: 3324
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030165924A1 197301.4
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 254, 378
;   OTHER INFORMATION: a, t, c, g, or other
US-10-240-965-168
```

```
  Query Match           23.0%;  Score 247;  DB 13;  Length 3324;
  Best Local Similarity   61.1%;  Pred. No. 2.3e-55;
  Matches  400;  Conservative    0;  Mismatches  255;  Indels    0;  Gaps    0;

Qy        290 AGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGC 349

Db        635 AGGTGTTTCCCATTAAACACTACCAGGACAAAATCAGGCCTTACATCCAGCTGCCATCGC 694

Qy        350 ACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTT 409

Db        695 ACAGCAACATTACTGGCATTGTGGAAGTGATCCTTGGGGAAACCAAGGCCTATGTCTTCT 754

Qy        410 TCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGC 469

Db        755 TTGAGAAGGACTTTGGGGACATGCACTCCTATGTGCGAAGCCGGAAGAGGCTGCGGGAAG 814

Qy        470 CTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACG 529

Db        815 AGGAAGCCGCCCGGCTCTTCAAGCAGATTGTCTCCGCCGTCGCCCACTGCCACCAGTCAG 874

Qy        530 GTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGA 589

Db        875 CCATCGTGCTGGGGGACCTGAAGCTTAGGAAGTTCGTCTTCTCCACGGAGGAGAGAACCC 934

Qy        590 AGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT 649

Db        935 AGCTTAGACTAGAAAGTCTAGAAGACACACACATAATGAAGGGGGAAGATGATGCTTTGT 994

Qy        650 GGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCAT 709

Db        995 CAGACAAACATGGCTGCCCAGCCTACGTGAGCCCTGAGATCCTCAACACCACTGGGACCT 1054

Qy        710 ACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCG 769

Db       1055 ACTCCGGAAAGGCTGCGGACGTTTGGAGCCTGGGGGTGATGCTCTACACCCTTCTGGTTG 1114

Qy        770 GCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGG 829

Db       1115 GACGATACCCCTTCCATGACTCAGACCCCAGTGCCCTTTTCTCCAAAATTCGGCGTGGAC 1174

Qy        830 CCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTC 889

Db       1175 AGTTCTGCATTCCTGAGCACATTTCCCCCAAAGCCAGGTGCCTCATTCGCAGCCTCTTGA 1234

Qy        890 GTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCT 944

Db       1235 GACGGGAGCCCTCCGAGAGACTCACTGCCCCCGAGATCCTACTGCACCCCTGGTT 1289


RESULT 12
US-10-228-263-1
; Sequence 1, Application US/10228263
; Publication No. US20030099985A1
; GENERAL INFORMATION:
;  APPLICANT: LI, Jing
;  TITLE OF INVENTION: AMPLIFIED GENE INVOLVED IN CANCER
;  FILE REFERENCE: 38002-0034
;  CURRENT APPLICATION NUMBER: US/10/228,263
;  CURRENT FILING DATE:  2002-12-04
;  PRIOR APPLICATION NUMBER: US 60/330,797
;  PRIOR FILING DATE: 2001-10-31
;  PRIOR APPLICATION NUMBER: US 60/314,655
;  PRIOR FILING DATE: 2001-08-27
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 818
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (788)..(788)
;   OTHER INFORMATION: N can be C or T
```

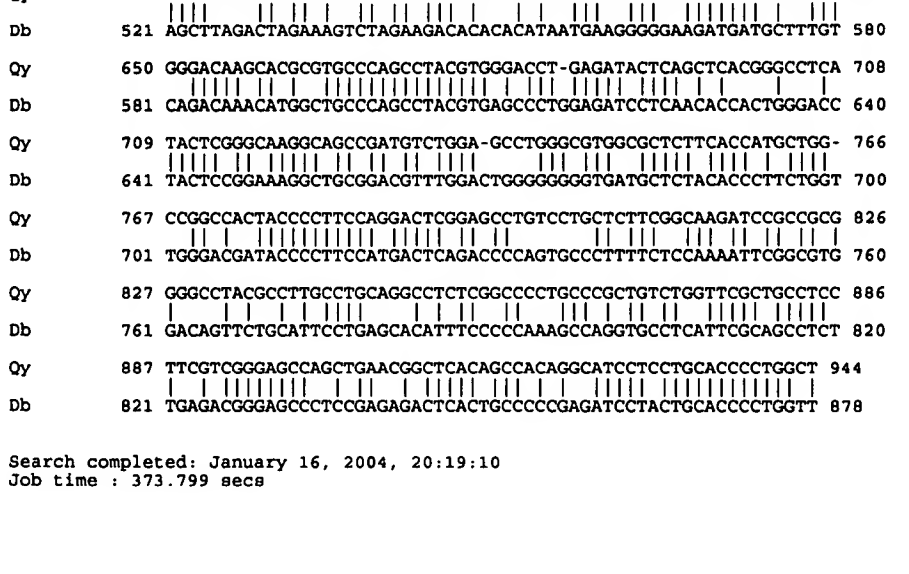
```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (56)..(676)
;   OTHER INFORMATION:
US-10-228-263-1

  Query Match             21.6%;  Score 232.4;  DB 15;  Length 818;
  Best Local Similarity   64.1%;  Pred. No. 1.5e-51;
  Matches  350;  Conservative    0;  Mismatches  196;  Indels    0;  Gaps    0;

Qy        399 CTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCG 458
Db         25 CTATGTCTTCTTTGAGAAGGACTTTGGGGACATGCACTCCTATGTGCGAAGCCGGAAGAG 84

Qy        459 TATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTG 518
Db         85 GCTGCGGGAAGAGGAAGCCGCCCGGCTCTTCAAGCAGATTGTCTCCGCCGTCGCCCACTG 144

Qy        519 TCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCG 578
Db        145 CCACCAGTCAGCCATCGTGCTGGGGGACCTGAAGCTTAGGAAGTTCGTCTTCTCCACGGA 204

Qy        579 TGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGA 638
Db        205 GGAGAGAACCCAGCTTAGACTAGAAAGTCTAGAAGACACACACATAATGAAGGGGGAAGA 264

Qy        639 TGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGACCTGAGATACTCAGCTC 698
Db        265 TGATGCTTTGTCAGACAAACATGGCTGCCCAGCCTACGTGAGCCCTGAGATCCTCAACAC 324

Qy        699 ACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCAC 758
Db        325 CACTGGGACCTACTCCGGAAAGGCTGCGGACGTTTGGAGCCTGGGGGTGATGCTCTACAC 384

Qy        759 CATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGAT 818
Db        385 CCTTCTGGTTGGACGATACCCCTTCCATGACTCAGACCCCAGTGCCCTTTTCTCCAAAAT 444

Qy        819 CCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCG 878
Db        445 TCGGCGTGGACAGTTCTGCATTCCTGAGCACATTTCCCCAAAGCCAGGTGCCTCATTCG 504

Qy        879 CTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCC 938
Db        505 CAGCCTCTTGAGACGGGAGCCCTCCGAGAGACTCACTGCCCCCGAGATCCTACTGCACCC 564

Qy        939 CTGGCT 944
Db        565 CTGGTT 570

RESULT 13
US-10-228-263-3
; Sequence 3, Application US/10228263
; Publication No. US20030099985A1
; GENERAL INFORMATION:
;  APPLICANT: LI, Jing
;  TITLE OF INVENTION: AMPLIFIED GENE INVOLVED IN CANCER
;  FILE REFERENCE: 38002-0034
;  CURRENT APPLICATION NUMBER: US/10/228,263
;  CURRENT FILING DATE:  2002-12-04
;  PRIOR APPLICATION NUMBER: US 60/330,797
;  PRIOR FILING DATE: 2001-10-31
;  PRIOR APPLICATION NUMBER: US 60/314,655
;  PRIOR FILING DATE: 2001-08-27
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-228-263-3

  Query Match             20.7%;  Score 222.2;  DB 15;  Length 621;
  Best Local Similarity   64.5%;  Pred. No. 7.2e-49;
  Matches  332;  Conservative    0;  Mismatches  183;  Indels    0;  Gaps    0;

Qy        430 ATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTC 489
Db          1 ATGCACTCCTATGTGCGAAGCCGGAAGAGGCTGCGGGAAGAGGAAGCCGCCCGGCTCTTC 60

Qy        490 CGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTC 549
Db         61 AAGCAGATTGTCTCCGCCGTCGCCCACTGCCACCAGTCAGCCATCGTGCTGGGGGACCTG 120

Qy        550 AAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTG 609
Db        121 AAGCTTAGGAAGTTCGTCTTCTCCACGGAGGAGAGAACCCAGCTTAGACTAGAAAGTCTA 180

Qy        610 GAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCA 669
Db        181 GAAGACACACACATAATGAAGGGGGAAGATGATGCTTTGTCAGACAAACATGGCTGCCCA 240

Qy        670 GCCTACGTGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGAT 729
Db        241 GCCTACGTGAGCCCTGAGATCCTCAACACCACTGGGACCTACTCCGGAAAGGCTGCGGAC 300

Qy        730 GTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGAC 789
Db        301 GTTTGGAGCCTGGGGGTGATGCTCTACACCCTTCTGGTTGGACGATACCCCTTCCATGAC 360

Qy        790 TCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGC 849
Db        361 TCAGACCCCAGTGCCCTTTTCTCCAAAATTCGGCGTGGACAGTTCTGCATTCCTGAGCAC 420

Qy        850 CTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGG 909
Db        421 ATTTCCCCAAAGCCAGGTGCCTCATTCGCAGCCTCTTGAGACGGGAGCCCTCCGAGAGA 480

Qy        910 CTCACAGCCACAGGCATCCTCCTGCACCCCTGGCT 944
Db        481 CTCACTGCCCCCGAGATCCTACTGCACCCCTGGTT 515

RESULT 14
US-09-925-301-525
; Sequence 525, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 525
;   LENGTH: 562
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (515)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (526)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (557)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-525
```

Query Match 20.6%; Score 220.8; DB 9; Length 562;
Best Local Similarity 96.1%; Pred. No. 1.7e-48;
Matches 219; Conservative 5; Mismatches 4; Indels 0; Gaps 0;



RESULT 15
US-10-291-172-492
; Sequence 492, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 492
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-291-172-492

Query Match 19.3%; Score 207.6; DB 12; Length 1943;
Best Local Similarity 60.2%; Pred. No. 5.2e-45;
Matches 396; Conservative 0; Mismatches 259; Indels 3; Gaps 3;



Search completed: January 16, 2004, 20:19:10
Job time : 373.799 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 10:41:02 ; Search time 68.5605 Seconds
(without alignments)
6914.267 Million cell updates/sec

Title: US-09-909-474D-1_COPY_49_1122
Perfect score: 1074
Sequence: 1 atgcgagccacccctctggc.....gagaagtggttctgtatggc 1074

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	956.4	89.1	1085	4	US-09-509-902A-3	Sequence 3, Appli
c 2	426.2	39.7	2559	4	US-09-220-132-135	Sequence 135, App
c 3	426.2	39.7	2562	1	US-08-146-421-4	Sequence 4, Appli
4	85	7.9	1302	1	US-08-913-050A-2	Sequence 2, Appli
5	85	7.9	1302	4	US-09-016-434-1146	Sequence 1146, Ap
6	81.6	7.6	1466	2	US-08-749-902-2	Sequence 2, Appli
7	71	6.6	288	4	US-09-016-434-391	Sequence 391, App
8	65.6	6.1	2169	4	US-09-016-434-1147	Sequence 1147, Ap
9	54.6	5.1	2902	4	US-09-579-664B-4	Sequence 4, Appli
c 10	51.2	4.8	30001	1	US-08-125-468-1	Sequence 1, Appli
c 11	51.2	4.8	30001	2	US-08-474-933-1	Sequence 1, Appli
12	51	4.7	601	4	US-09-691-861A-14	Sequence 14, Appl
13	50.4	4.7	1104	4	US-09-691-861A-1	Sequence 1, Appli
14	50.4	4.7	9862	4	US-09-691-861A-3	Sequence 3, Appli
15	49.4	4.6	263	4	US-09-016-434-536	Sequence 536, App
16	48.4	4.5	2598	4	US-09-816-094-1	Sequence 1, Appli
17	47.4	4.4	1349	1	US-07-951-715A-20	Sequence 20, Appl
18	47.4	4.4	1349	2	US-08-459-448A-20	Sequence 20, Appl
19	47.4	4.4	1349	3	US-08-459-595A-20	Sequence 20, Appl
20	47.4	4.4	1349	3	US-08-459-504B-20	Sequence 20, Appl
21	47.4	4.4	1349	3	US-08-459-444-20	Sequence 0, Appli
22	47.4	4.4	1349	4	US-09-547-422-20	Sequence 0, Appli
23	47	4.4	1636	4	US-09-016-434-1433	Sequence 1433, Ap
24	46.8	4.4	1742	3	US-08-557-006C-38	Sequence 38, Appl
25	46.8	4.4	2652	3	US-08-557-006C-39	Sequence 39, Appl
26	46.8	4.4	2761	3	US-08-557-006C-24	Sequence 24, Appl
27	46	4.3	7218	1	US-08-232-463-14	Sequence 14, Appl
28	45.8	4.3	7301	4	US-09-816-094-3	Sequence 3, Appli
29	45.2	4.2	1599	3	US-09-256-465-1	Sequence 1, Appli
30	45.2	4.2	1599	4	US-09-167-322-3	Sequence 3, Appli
31	44.2	4.1	1282	2	US-08-878-989-12	Sequence 12, Appl
32	44.2	4.1	1282	3	US-09-272-796-12	Sequence 12, Appl
33	44.2	4.1	1282	4	US-09-016-434-953	Sequence 953, App
34	43.2	4.0	4162	2	US-08-459-448A-26	Sequence 26, Appl
35	43.2	4.0	4162	3	US-08-459-595A-26	Sequence 26, Appl
36	43.2	4.0	4162	3	US-08-459-504B-26	Sequence 26, Appl
37	43.2	4.0	4162	3	US-08-459-444-26	Sequence 26, Appl
38	43.2	4.0	4162	4	US-09-547-422-26	Sequence 26, Appl
39	43.2	4.0	4165	1	US-07-951-715A-26	Sequence 26, Appl
40	42.4	3.9	1879	2	US-08-403-852D-5	Sequence 5, Appli
41	42.4	3.9	1879	3	US-08-510-646B-5	Sequence 5, Appli
42	42.4	3.9	1879	3	US-09-231-818-5	Sequence 5, Appli
43	42.4	3.9	15872	3	US-09-105-537-1	Sequence 1, Appli
44	42.2	3.9	759	4	US-09-769-482-2	Sequence 2, Appli
45	42.2	3.9	1140	4	US-09-769-482-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-509-902A-3
; Sequence 3, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1085
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-3

Query Match 89.1%; Score 956.4; DB 4; Length 1085;
Best Local Similarity 99.9%; Pred. No. 3.2e-207;
Matches 957; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 187

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 247

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 307

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 367

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 427

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 487

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
```

```
Db        488 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 547

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Db        548 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 607

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Db        608 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 667

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Db        668 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 727

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Db        728 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 787

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Db        788 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 847

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Db        848 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 907

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840

Db        908 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 967

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900

Db        968 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1027

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 958

Db       1028 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1085
```

```
RESULT 2
US-09-220-132-135/c
; Sequence 135, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  METHODS AND COMPOSITONS FOR THE IDENTIFICATION AND ASSESSMENT
;  TITLE OF INVENTION:  OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCE
;  FILE REFERENCE:  07334-074001
;  CURRENT APPLICATION NUMBER: US/09/220,132
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  US 60/079,303
;  PRIOR FILING DATE:  1998-03-25
;  PRIOR APPLICATION NUMBER:  US 60/068,821
;  PRIOR FILING DATE:  1997-12-24
;  NUMBER OF SEQ ID NOS:  191
;  SOFTWARE:  PastSEQ for Windows Version 4.0
; SEQ ID NO 135
;   LENGTH: 2559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-220-132-135

  Query Match             39.7%;  Score 426.2;  DB 4;  Length 2559;
  Best Local Similarity   93.8%;  Pred. No. 1.7e-87;
  Matches  487;  Conservative    0;  Mismatches   28;  Indels    4;  Gaps    4;

Qy        556 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 615

Db       1634 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 1575

Qy        616 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 675

Db       1574 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT-GGACAAGCACGCGTCGCCAGCCTAC 1516

Qy        676 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 735

Db       1515 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 1456

Qy        736 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 795

Db       1455 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 1396

Qy        796 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 855

Db       1395 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGCGCTCTCG 1336

Qy        856 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 915

Db       1335 CCCTCGCCGCGCTGTCTGGTTCCGTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 1276

Qy        916 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 975

Db       1275 G-CACAGCGATCCTCCTGCACCCCTGGCTGCGACA-GACCCGATGCCCTTAGCCCCAACC 1218

Qy        976 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1035

Db       1217 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA-CC 1159

Qy       1036 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074

Db       1158 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1120
```

```
RESULT 3
US-08-146-421-4/c
; Sequence 4, Application US/08146421
; Patent No. 5543499
;  GENERAL INFORMATION:
;    APPLICANT:  BREWER, GARY
;    TITLE OF INVENTION:  DNA SEQUENCE ENCODING A POLYPEPTIDE WITH
;    TITLE OF INVENTION:  ANTI-TUMOR PROPERTIES
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DILWORTH & BARRESE
;      STREET:  4350 LA JOLLA VILLAGE DRIVE, SUITE 300
;      CITY:  SAN DIEGO
;      STATE:  CALIFORNIA
;      COUNTRY:  U.S.A.
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/146,421
;      FILING DATE:  29-OCT-1993
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PEPPER PH.D., FREDERICK W.
;      REGISTRATION NUMBER:  31,286
;      REFERENCE/DOCKET NUMBER:  489-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-4410
;      TELEFAX:  619-453-2839
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2562 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  246..1106
```

```
US-08-146-421-4

  Query Match             39.7%;  Score 426.2;  DB 1;  Length 2562;
  Best Local Similarity   93.8%;  Pred. No. 1.7e-87;
  Matches  487;  Conservative    0;  Mismatches   28;  Indels    4;  Gaps    4;

Qy        556 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 615

Db       1637 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 1578

Qy        616 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGACAAGCACGCGTGCCCAGCCTAC 675

Db       1577 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT-GGACAAGCACGCGTCGCCAGCCTAC 1519

Qy        676 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1518 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 1459

Qy        736 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1458 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 1399

Qy        796 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 855

Db       1398 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGCGCTCTCG 1339

Qy        856 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 915

Db       1338 CCCTCGCCGCGCTGTCTGGTTCCGTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 1279

Qy        916 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 975

Db       1278 G-CACAGCGATCCTCCTGCACCCCTGGCTGCGACA-GACCCGATGCCCTTAGCCCCAACC 1221

Qy        976 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1035

Db       1220 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA-CC 1162

Qy       1036 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              |||||||||||||||||||||||||||||||||||||||
Db       1161 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1123

RESULT 4
US-08-913-050A-2
; Sequence 2, Application US/08913050A
; Patent No. 5827726
;  GENERAL INFORMATION:
;    APPLICANT:  NEZU, Jun-ichi
;    TITLE OF INVENTION:  DNA ENCODING PROTEIN KINASE
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK, P.L.L.C.
;      STREET:  419 7th Street N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/913,050A
;      FILING DATE:  05-SEP-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 57104/1995
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP PCT/JP96/00660
;      FILING DATE:  15-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YUN, Allen C.
;      REGISTRATION NUMBER:  37,971
;      REFERENCE/DOCKET NUMBER:  NEZU=4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1302 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-913-050A-2

  Query Match             7.9%;  Score 85;  DB 1;  Length 1302;
  Best Local Similarity   48.1%;  Pred. No. 1.5e-10;
  Matches  241;  Conservative    0;  Mismatches  260;  Indels    0;  Gaps    0;

Qy        455 ACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGC 514

Db        437 AGCGTTTCCCAGTGTGCCAGGCCCACGGGTACTTCTGTCAGCTGATTGACGGCCTGGAGT 496

Qy        515 ACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTG 574

Db        497 ACCTGCATAGCCAGGGCATTGTGCACAAGGACATCAAGCCGGGGAACCTGCTGCTCACCA 556

Qy        575 ACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 634

Db        557 CCGGTGGCACCCTCAAAATCTCCGACCTGGGCGTGGCCGAGGCACTGCACCCGTTCGCGG 616

Qy        635 CAGATGATTCCCTGTGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCA 694

Db        617 CGGACGACACCTGCCGGACCAGCCAGGGCTCCCCGGCTTTCCAGCCGCCCGAGATTGCCA 676

Qy        695 GCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCT 754

Db        677 ACGGCCTGGACACCTTCTCCGGCTTCAAGGTGGACATCTGGTCGGCTGGGGTCACCCTCT 736

Qy        755 TCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCA 814

Db        737 ACAACATCACCACGGGTCTGTACCCCTTCGAAGGGGACAACATCTACAAGTTGTTTGAGA 796

Qy        815 AGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGG 874

Db        797 ACATCGGGAAGGGGAGCTACGCCATCCCGGGCGACTGTGGCCCCCCGCTCTCTGACCTGC 856

Qy        875 TTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGC 934

Db        857 TGAAAGGGATGCTTGAGTACGAACCGGCCAAGAGGTTCTCCATCCGGCAGATCCGGCAGC 916

Qy        935 ACCCCTGGCTGCGACAGGACC 955

Db        917 ACAGCTGGTTCCGGAAGAAAC 937

RESULT 5
US-09-016-434-1146
; Sequence 1146, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1146:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1302 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g1480860
US-09-016-434-1146

  Query Match             7.9%;  Score 85;  DB 4;  Length 1302;
  Best Local Similarity   48.1%;  Pred. No. 1.5e-10;
  Matches  241;  Conservative    0;  Mismatches  260;  Indels    0;  Gaps    0;

Qy        455 ACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGC 514
              | ||| |||| | |      ||||    || |  |||| | ||| ||      || ||||| |
Db        437 AGCGTTTCCCAGTGTGCCAGGCCCACGGGTACTTCTGTCAGCTGATTGACGGCCTGGAGT 496

Qy        515 ACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTG 574
               ||   ||     || ||  | || |      ||  |||||| | |     |  |   | ||  |
Db        497 ACCTGCATAGCCAGGGCATTGTGCACAAGGACATCAAGCCGGGGAACCTGCTGCTCACCA 556

Qy        575 ACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 634
                | |||  |       ||  |            |  | |||  ||  |        | |    | |
Db        557 CCGGTGGCACCCTCAAAATCTCCGACCTGGGCGTGGCCGAGGCACTGCACCCGTTCGCGG 616

Qy        635 CAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCA 694
              | || || ||     ||  ||   || |  | ||| || |  |     || ||||     ||
Db        617 CGGACGACACCTGCCGGACCAGCCAGGGCTCCCCGGCTTTCCAGCCGCCCGAGATTGCCA 676

Qy        695 GCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCT 754
                |   | || | | | ||| |||       |  ||  |||||         || ||   | ||||
Db        677 ACGGCCTGGACACCTTCTCCGGCTTCAAGGTGGACATCTGGTCGGCTGGGGTCACCCTCT 736

Qy        755 TCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCA 814
                || |||      ||| |  ||||||||| | |       |       |  | | || |  |
Db        737 ACAACATCACCACGGGTCTGTACCCCTTCGAAGGGGACAACATCTACAAGTTGTTTGAGA 796

Qy        815 AGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGG 874
              | ||| |     |||  ||||||| | || |  |  |        ||||   |        |||
Db        797 ACATCGGGAAGGGGAGCTACGCCATCCCGGGCGACTGTGGCCCCCCGCTCTCTGACCTGC 856

Qy        875 TTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGC 934
              |       |  | |||          || || ||  |   | || || |    |       |||| |  ||
Db        857 TGAAAGGGATGCTTGAGTACGAACCGGCCAAGAGGTTCTCCATCCGGCAGATCCGGCAGC 916

Qy        935 ACCCCTGGCTGCGACAGGACC 955
               ||  |||| | ||   || | |
Db        917 ACAGCTGGTTCCGGAAGAAAC 937
```

RESULT 6

```
US-08-749-902-2
; Sequence 2, Application US/08749902
; Patent No. 5985635
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    TITLE OF INVENTION:  NOVEL HUMAN SERINE/THREONINE
;    TITLE OF INVENTION:  PROTEIN KINASES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/749,902
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0150 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1466 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  Consensus
US-08-749-902-2

  Query Match             7.6%;  Score 81.6;  DB 2;  Length 1466;
  Best Local Similarity   47.3%;  Pred. No. 9.1e-10;
  Matches  237;  Conservative    0;  Mismatches  264;  Indels    0;  Gaps    0;

Qy        455 ACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGC 514
              | ||| |||| | |      ||||    || |  |||| | ||| ||      || | ||| |
Db        536 AGCGTTTCCCAGTGTGCCAGGCCCACGGGTACTTCTGTCAGCTGATTGACGGCNTGGGT 595

Qy        515 ACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTG 574
               ||   ||     || ||  | || |      |    |||||| | |     |  |   | ||  |
Db        596 ACCTGCATNGCCAGGGNATTGTGCACAAGGGCATCAAGCCGGGGAACCTGCTGCTCACCA 655

Qy        575 ACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 634
                | |||  |       ||  |            |  | |||  ||  |        | |    |
Db        656 CCGGTGGCACCCTCAAAATCTCCGACCTGGGCGTGGCCGAGGCACTGCACCCGTTCGCNG 715

Qy        635 CAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCA 694
              | || ||  ||  ||   || |  | ||| || |  |     || || |     ||
Db        716 CGGACGACACCTGCCGGACCAGCCAGGGCTCCCCGGCTTTCCAGCCGCCCGANATTGCCA 775

Qy        695 GCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCT 754
                |   | || | | | ||| |||       |  ||  |||||         || ||   | ||||
Db        776 ACGGCCTGGACACCTTCTCCGGCTTCAAGGTGGACATCTGGTCGGCTGGGGTCACCCTCT 835

Qy        755 TCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCA 814
                || |||      ||| |  ||||||||| | |       |       |  | | || |  |
Db        836 ACAACATCACCACGGGTCTGTACCCCTTCGAAGGGGACAACATCTACAAGTTGTTTGAGA 895
```

```
Qy        815 AGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGG 874
Db        896 ACATCGGGAAGGGGAGCTACGCCATCCCGGGCGACTGTGGCCCCCCGCTCTCTGACCTGC 955

Qy        875 TTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGC 934
Db        956 TGAAAGGGATGCTTGAGTACGAACCGGCCAAGAGGTTCTCCATCCGGCAGATCCGGCAGC 1015

Qy        935 ACCCCTGGCTGCGACAGGACC 955
Db       1016 ACAGCTGGTTCCGGAAGAAAC 1036

RESULT 7
US-09-016-434-391
; Sequence 391, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  391:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  288 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRAITUT03
;      CLONE:  2108752
US-09-016-434-391

  Query Match             6.6%;  Score 71;  DB 4;  Length 288;
  Best Local Similarity   63.3%;  Pred. No. 1.7e-07;
  Matches  126;  Conservative    0;  Mismatches   70;  Indels    3;  Gaps    1;

Qy        646 CTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCC 705
Db         69 CTGAAGGACCAGAGAGGGAGCCCTGCCTACATCAGTCCCGACGTGCTCAGCGGCCGG--- 125

Qy        706 TCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTG 765
Db        126 CCGTACCGTGGCAAGCCCAGTGACATGTGGGCCCTGGGCGTGGTGCTCTTCACCATGCTG 185

Qy        766 GCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGC 825
Db        186 TATGGCCAGTTCCCCTTCTACGACAGCATCCCGCAGGAGCTCTTCCGCAAGATCAAGGCT 245

Qy        826 GGGGCCTACGCCTTGCCTG 844
Db        246 GCCGAGTATACCATTCCTG 264

RESULT 8
US-09-016-434-1147
; Sequence 1147, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1147:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2169 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g1488262
US-09-016-434-1147

  Query Match             6.1%;  Score 65.6;  DB 4;  Length 2169;
  Best Local Similarity   45.9%;  Pred. No. 4e-06;
  Matches  310;  Conservative    0;  Mismatches  354;  Indels   12;  Gaps    2;

Qy        288 CAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCC 347
Db        216 CAAGCCGCATCAGCGCGAGAAGATCCTAAATGAGATTGAGCTGCACCGAGACCTGCAGCA 275

Qy        348 GCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTT 407
Db        276 CCGCCACATCGTGCGTTTTTCGCACCACTTTGAGGACGCTGACAACATCTACATTTTCTT 335

Qy        408 TTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGA 467
Db        336 GGAGCTCTGCAGCCGAAAGTCCCTGGCCCACATCTGGAAGGCCCGGCACACCCTGTTGGA 395
```

```
Qy        468 GCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCA 527
Db        396 GCCAGAAGTGCGCTACTACCTGCGGCAGATCCTTTCTGGCCTCAAGTACTTGCACCAGCG 455

Qy        528 CGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAA 587
Db        456 CGGCATCTTGCACCGGGACCTCAAGTTGGGAAATTTTTTCATCAC------TGAGAACAT 509

Qy        588 GAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCT 647
Db        510 GGAACTGAAGGTGGGGGATTTTGGGCTGGCAGCCCGGTTGGAGCCTCCGGAGCAGAGGAA 569

Qy        648 GTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTC 707
Db        570 GAAGACCATCTGTGGCACCCCCAACTATGTGGCTCCAGAAGTGCTGCTGAGACAGGGCCA 629

Qy        708 ATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGC 767
Db        630 C------GGCCCTGAAGCGGATGTATGGTCACTGGGCTGTGTCATGTACACGCTGCTCTG 683

Qy        768 CGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGG 827
Db        684 CGGGAGCCCTCCCTTTGAGACGGCTGACCTGAAGGAGACGTACCGCTGCATCAAGCAGGT 743

Qy        828 GGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCT 887
Db        744 TCACTACACGCTGCCTGCCAGCCTCTCACTGCCTGCCCGGCAGCTCCTGGCCGCCATCCT 803

Qy        888 TCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCG 947
Db        804 TCGGGCCTCACCCCGAGACCGCCCCTCTATTGACCAGATCCTGCGCCATGACTTCTTTAC 863

Qy        948 ACAGGACCCGATGCCC 963
Db        864 CAAGGGCTACACCCCC 879
```

RESULT 9
US-09-579-664B-4
; Sequence 4, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-579-664B-4

Query Match 5.1%; Score 54.6; DB 4; Length 2902;
Best Local Similarity 49.7%; Pred. No. 0.0013;
Matches 231; Conservative 0; Mismatches 219; Indels 15; Gaps 3;

```
Qy        451 CGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTG 510
Db        564 CGGCCACGGCTGAGTGAGCGGGACGCCAGGCATTTCTTCCGACAGATCGTGTCTGCCCTG 623

Qy        511 GCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTC 570
Db        624 CACTACTGCCACCAGAACGGGATCGTTCACCGAGATCTCAAGCTGGAAAACATCCTTCTA 683

Qy        571 GCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACT 630
Db        684 GATGCCAATGGAAACATCAAGATTGCTGACTTTGGCCT---------CTCCAACCTGTAC 734

Qy        631 GGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATA 690
Db        735 CACAAAGGCAAGTTCCTCCAGACGTTCTGTGGGAGCCCTCTCTACGCCTCGCCTGAGATA 794

Qy        691 CTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCG 750
Db        795 GTCAACGGGAAGCCC---TATGTGGGCCCAGAGGTGGACAGCTGGTCTCTGGGCGTTCTC 851

Qy        751 CTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTC 810
Db        852 CTGTACATCCTGGTGCATGGCACCATGCCCTTTGACGGGCAGGATCATAAAACACTGGTG 911

Qy        811 GGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGT 870
Db        912 AAGCAAATCAGTAACGGGGCTTACCGTGAGCCGCC---CAAGCCGTCCGATGCCTGTGGC 968

Qy        871 CTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 915
Db        969 CTGATCCGGTGGCTGTTAATGGTGAACCCCACCCGTCGGGCCACA 1013
```

RESULT 10
US-08-125-468-1/c
; Sequence 1, Application US/08125468
; Patent No. 5589385
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fantini, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlortetracycline and tetracyline Formation and cosmids
; TITLE OF INVENTION: useful therein
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,468
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,255-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)831-3241
; TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-125-468-1

Query Match 4.8%; Score 51.2; DB 1; Length 30001;
Best Local Similarity 44.2%; Pred. No. 0.011;
Matches 212; Conservative 0; Mismatches 268; Indels 0; Gaps 0;

```
Qy        390 CCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAG 449
```

```
Db      22561 CGAGAGGCTCAACGGGATGTTCGCCTTCGCCGTCTGGGACGAGCACAGCGAGGAACTCCT 22502

Qy        450 CCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCT 509

Db      22501 GCTGGTCCGCGACCGGATGGGCGTCAAGCCGCTCTACTACTGGCCGACCGCCGCCGGCGT 22442

Qy        510 GGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTT 569

Db      22441 GCGCTTCGGCTCGGAGCCCAAGGCCGTCCTGGCGGACCGGACCCTGTCCCGCCGGGTCGG 22382

Qy        570 CGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGAC 629

Db      22381 CCCGGACGGGCTGTGCGAGGTGCTGGACATGGTGAAGACGCCCGAGGCGGCCGTCTTCTC 22322

Qy        630 TGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGAT 689

Db      22321 CGGGCTCGTCGAGGTCCGTCCGGGTCAGCTCGTCCGGGTGGGGCGCGGCGGGGTGCGCCG 22262

Qy        690 ACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGC 749

Db      22261 GCAGACGTACTGGCGGCTGGAGGCCCGCGAGCACACCGACGACCTGCCGACGACCATCGG 22202

Qy        750 GCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTT 809

Db      22201 CACCGTCCGTGACCTGCTCGCCGACACCGTCACCCGTCAGCTGGTCTCGGACGTCCCGCT 22142

Qy        810 CGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTG 869

Db      22141 GGGCACGCTGCTCTCCGGCGGCCTGGACTCGTCGGCCGTGACGGCGCTGGCCGCCCGCAG 22082


RESULT 11
US-08-474-933-1/c
; Sequence 1, Application US/08474933
; Patent No. 5866410
;  GENERAL INFORMATION:
;    APPLICANT:  Ryan, Michael J.
;    APPLICANT:  Lotvin, Jason A.
;    APPLICANT:  Strathy, Nancy
;    APPLICANT:  Fantini, Susan E.
;    TITLE OF INVENTION:  Cloning of the biosynthetic pathway for
;    TITLE OF INVENTION:  chlortetracycline and tetracyline Formation and cosmids
;    TITLE OF INVENTION:  useful therein
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  American Cyanamid Company
;      STREET:  One Cyanamid Plaza
;      CITY:  Wayne
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07470
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,933
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/125,468
;      FILING DATE:  22-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tsevdos, Estelle J
;      REGISTRATION NUMBER:  31,145
;      REFERENCE/DOCKET NUMBER:  31,255-02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201)831-3241
;      TELEFAX:  (201)831-3305
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-474-933-1

  Query Match             4.8%;  Score 51.2;  DB 2;  Length 30001;
  Best Local Similarity   44.2%;  Pred. No. 0.011;
  Matches  212;  Conservative    0;  Mismatches  268;  Indels    0;  Gaps    0;

Qy        390 CCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAG 449

Db      22561 CGAGAGGCTCAACGGGATGTTCGCCTTCGCCGTCTGGGACGAGCACAGCGAGGAACTCCT 22502

Qy        450 CCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCT 509

Db      22501 GCTGGTCCGCGACCGGATGGGCGTCAAGCCGCTCTACTACTGGCCGACCGCCGCCGGCGT 22442

Qy        510 GGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTT 569

Db      22441 GCGCTTCGGCTCGGAGCCCAAGGCCGTCCTGGCGGACCGGACCCTGTCCCGCCGGGTCGG 22382

Qy        570 CGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGAC 629

Db      22381 CCCGGACGGGCTGTGCGAGGTGCTGGACATGGTGAAGACGCCCGAGGCGGCCGTCTTCTC 22322

Qy        630 TGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGAT 689

Db      22321 CGGGCTCGTCGAGGTCCGTCCGGGTCAGCTCGTCCGGGTGGGGCGCGGCGGGGTGCGCCG 22262

Qy        690 ACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGC 749

Db      22261 GCAGACGTACTGGCGGCTGGAGGCCCGCGAGCACACCGACGACCTGCCGACGACCATCGG 22202

Qy        750 GCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTT 809

Db      22201 CACCGTCCGTGACCTGCTCGCCGACACCGTCACCCGTCAGCTGGTCTCGGACGTCCCGCT 22142

Qy        810 CGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTG 869

Db      22141 GGGCACGCTGCTCTCCGGCGGCCTGGACTCGTCGGCCGTGACGGCGCTGGCCGCCCGCAG 22082


RESULT 12
US-09-691-861A-14
; Sequence 14, Application US/09691861A
; Patent No. 6482935
; GENERAL INFORMATION:
;  APPLICANT: Wei, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL000892
;  CURRENT APPLICATION NUMBER: US/09/691,861A
;  CURRENT FILING DATE:  2000-10-18
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(601)
;   OTHER INFORMATION: n = A,T,C or G
US-09-691-861A-14

  Query Match             4.7%;  Score 51;  DB 4;  Length 601;
  Best Local Similarity   49.7%;  Pred. No. 0.0062;
  Matches  188;  Conservative    0;  Mismatches  181;  Indels    9;  Gaps    2;

Qy        420 CCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGC 479
```

```
Db        73 CCAGGGCGACCTCCTCGAGTTAATCAAAACCCGGGGAGCCCTGCATGAGGACGAAGCTCG 132

Qy       480 CGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCT 539

Db       133 CAAGAAGTTCCACCAGCTTTCCTTGGCCATCAAGTACTGCCACGACCTGGACGTCGTCCA 192

Qy       540 GCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCT 599

Db       193 CCGGGACCTCAAGTGTGACAACCTTCTCCTTGACAAGGACTTCAACATCAAGCTGTCCGA 252

Qy       600 GGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCA 659

Db       253 CTTCAGCTTCTCCAAGCGCTGCCTGCGGGATGACAGTGGTCGAATGGCNTTAAGCAAGAC 312

Qy       660 C------GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTC 713

Db       313 CTTCTGTGGGTCACCAGCGTATGCGGCCCCAGAGGTGCT---GCAGGGCATTCCCTACCA 369

Qy       714 GGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCA 773

Db       370 GCCCAAGGTGTACGACATCTGGAGCCTAGGCGTGATCCTCTACATCATGGTCTGCGGCTC 429

Qy       774 CTACCCCTTCCAGGACTC 791

Db       430 CATGCCCTACGACGACTC 447


RESULT 13
US-09-691-861A-1
; Sequence 1, Application US/09691861A
; Patent No. 6482935
; GENERAL INFORMATION:
;  APPLICANT: Wei, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL000892
;  CURRENT APPLICATION NUMBER: US/09/691,861A
;  CURRENT FILING DATE:  2000-10-18
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1104
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-691-861A-1

  Query Match             4.7%;  Score 50.4;  DB 4;  Length 1104;
  Best Local Similarity   49.7%;  Pred. No. 0.0095;
  Matches  188;  Conservative    0;  Mismatches  181;  Indels    9;  Gaps    2;

Qy       420 CCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGC 479

Db       282 CCAGGGCGACCTCCTCGAGTTAATCAAAACCCGGGGAGCCCTGCATGAGGACGAAGCTCG 341

Qy       480 CGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCT 539

Db       342 CAAGAAGTTCCACCAGCTTTCCTTGGCCATCAAGTACTGCCACGACCTGGACGTCGTCCA 401

Qy       540 GCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCT 599

Db       402 CCGGGACCTCAAGTGTGACAACCTTCTCCTTGACAAGGACTTCAACATCAAGCTGTCCGA 461

Qy       600 GGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCA 659

Db       462 CTTCAGCTTCTCCAAGCGCTGCCTGCGGGATGACAGTGGTCGAATGGCATTAAGCAAGAC 521

Qy       660 C------GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTC 713

Db       522 CTTCTGTGGGTCACCAGCGTATGCGGCCCCAGAGGTGCT---GCAGGGCATTCCCTACCA 578

Qy       714 GGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCA 773

Db       579 GCCCAAGGTGTACGACATCTGGAGCCTAGGCGTGATCCTCTACATCATGGTCTGCGGCTC 638

Qy       774 CTACCCCTTCCAGGACTC 791

Db       639 CATGCCCTACGACGACTC 656


RESULT 14
US-09-691-861A-3
; Sequence 3, Application US/09691861A
; Patent No. 6482935
; GENERAL INFORMATION:
;  APPLICANT: Wei, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL000892
;  CURRENT APPLICATION NUMBER: US/09/691,861A
;  CURRENT FILING DATE:  2000-10-18
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 9862
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-691-861A-3

  Query Match             4.7%;  Score 50.4;  DB 4;  Length 9862;
  Best Local Similarity   49.7%;  Pred. No. 0.014;
  Matches  188;  Conservative    0;  Mismatches  181;  Indels    9;  Gaps    2;

Qy       420 CCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGC 479

Db      1784 CCAGGGCGACCTCCTCGAGTTAATCAAAACCCGGGGAGCCCTGCATGAGGACGAAGCTCG 1843

Qy       480 CGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCT 539

Db      1844 CAAGAAGTTCCACCAGCTTTCCTTGGCCATCAAGTACTGCCACGACCTGGACGTCGTCCA 1903

Qy       540 GCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCT 599

Db      1904 CCGGGACCTCAAGTGTGACAACCTTCTCCTTGACAAGGACTTCAACATCAAGCTGTCCGA 1963

Qy       600 GGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCA 659

Db      1964 CTTCAGCTTCTCCAAGCGCTGCCTGCGGGATGACAGTGGTCGAATGGCCTTAAGCAAGAC 2023

Qy       660 C------GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTC 713

Db      2024 CTTCTGTGGGTCACCAGCGTATGCGGCCCCAGAGGTGCT---GCAGGGCATTCCCTACCA 2080

Qy       714 GGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCA 773

Db      2081 GCCCAAGGTGTACGACATCTGGAGCCTAGGCGTGATCCTCTACATCATGGTCTGCGGCTC 2140

Qy       774 CTACCCCTTCCAGGACTC 791

Db      2141 CATGCCCTACGACGACTC 2158


RESULT 15
US-09-016-434-536
; Sequence 536, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
```

```
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/016,434
;       FILING DATE:  HEREWITH
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Zeller, Karen J.
;       REGISTRATION NUMBER:  37,071
;       REFERENCE/DOCKET NUMBER:  PA-0002 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (650) 855-0555
;       TELEFAX:  (650) 845-4166
;   INFORMATION FOR SEQ ID NO:  536:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  263 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  NGANNOT01
;       CLONE:  2314392
US-09-016-434-536

  Query Match             4.6%;  Score 49.4;  DB 4;  Length 263;
  Best Local Similarity   47.3%;  Pred. No. 0.012;
  Matches  107;  Conservative    0;  Mismatches  119;  Indels    0;  Gaps    0;

Qy        727 GATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAG 786
              ||  |||||   ||||  |       | | |||  ||||| | |||  |   ||||||  |
Db         10 GACATCTGGNCTCTGGNCTGCATCATGTACACGGTGCTGACTGGCACCCCACCCTTCATG 69

Qy        787 GACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCA 846
              | |||    |       | | | |   || |||||    |    |||| ||   |  ||
Db         70 GCCTCACCCCTGNNGGAGATGTACCAAAACATCCGTTANGNCCACTACNCCGANCNCGCT 129

Qy        847 GGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAA 906
              || || | |  | | ||||| ||  |     ||||||    |     |  || ||
Db        130 CANCTGTCTGNCAATNCTCGCTGCCTCATNGTTCACCTCCTANCACCCANNCNNGCNGAT 189

Qy        907 CGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGG 952
              |  | ||            | || | | ||  ||   |   |||||
Db        190 CNNCCCAGCCTGGNCCANCTGCTGCAGGACGACTTCTTCATACAGG 235


Search completed: January 16, 2004, 16:49:45
Job time : 72.5605 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 10:41:02 ; Search time 131.44 Seconds
(without alignments)
6914.267 Million cell updates/sec

Title: US-09-909-474D-1
Perfect score: 2059
Sequence: 1 gctctgagccccggcggcgc.....aaaaaaaaaaaaaaaaaaa 2059

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1004.4	48.8	1085	4	US-09-509-902A-3	Sequence 3, Appli
c	2	505.2	24.5	2559	4	US-09-220-132-135	Sequence 135, App
c	3	505.2	24.5	2562	1	US-08-146-421-4	Sequence 4, Appli
	4	85	4.1	1302	1	US-08-913-050A-2	Sequence 2, Appli
	5	85	4.1	1302	4	US-09-016-434-1146	Sequence 1146, Ap
	6	84.6	4.1	1545	4	US-09-559-023-1	Sequence 1, Appli
	7	82.8	4.0	1674	4	US-09-996-243-300	Sequence 300, App
	8	82.4	4.0	2136	4	US-09-996-243-302	Sequence 302, App
	9	81.6	4.0	1466	2	US-08-749-902-2	Sequence 2, Appli
	10	79.8	3.9	1048	4	US-09-489-847-38	Sequence 38, Appl
	11	79.6	3.9	485	4	US-09-996-243-245	Sequence 245, App
	12	79.6	3.9	1066	1	US-08-157-101A-4	Sequence 4, Appli
	13	79.6	3.9	1927	4	US-09-336-536-66	Sequence 66, Appl
	14	79.2	3.8	1091	4	US-09-328-965-1	Sequence 1, Appli
	15	79.2	3.8	1882	3	US-09-370-253-1	Sequence 1, Appli
	16	78.8	3.8	1024	4	US-09-328-475C-50	Sequence 50, Appl
	17	78.8	3.8	2674	4	US-09-817-180-1	Sequence 1, Appli
	18	78.4	3.8	2246	4	US-09-363-708-3	Sequence 3, Appli
	19	78.4	3.8	2246	4	US-09-083-587-3	Sequence 3, Appli
	20	78.2	3.8	1342	4	US-09-489-847-89	Sequence 89, Appl
	21	78.2	3.8	2434	4	US-09-489-847-67	Sequence 67, Appl
	22	78	3.8	1582	3	US-08-545-196B-10	Sequence 10, Appl
	23	78	3.8	1582	3	US-08-545-196B-12	Sequence 12, Appl
	24	77.8	3.8	903	5	PCT-US95-06406A-21	Sequence 21, Appl
	25	77.8	3.8	1872	4	US-09-801-052-1	Sequence 1, Appli
	26	77.8	3.8	2323	4	US-09-149-476-24	Sequence 24, Appl
	27	77.8	3.8	3715	4	US-09-234-245-1	Sequence 1, Appli
	28	77.4	3.8	140	1	US-08-628-417-5	Sequence 5, Appli
	29	77.4	3.8	240	1	US-08-628-417-6	Sequence 6, Appli
	30	77.4	3.8	2184	3	US-08-955-918C-1	Sequence 1, Appli
	31	77.4	3.8	2184	4	US-08-697-766A-1	Sequence 1, Appli
	32	77.4	3.8	2186	4	US-09-360-545-66	Sequence 66, Appl
	33	77.4	3.8	3060	4	US-09-996-243-504	Sequence 504, App
	34	77.2	3.7	1249	4	US-09-461-325-128	Sequence 128, App
c	35	77.2	3.7	1260	4	US-09-461-325-93	Sequence 93, Appl
	36	77.2	3.7	2202	4	US-09-465-558-59	Sequence 59, Appl
	37	77.2	3.7	2447	2	US-09-014-969-14	Sequence 14, Appl
	38	77	3.7	1198	3	US-09-248-335-27	Sequence 27, Appl
	39	77	3.7	1414	4	US-09-501-115-5	Sequence 5, Appli
	40	76.8	3.7	1411	3	US-08-964-127-5	Sequence 5, Appli
	41	76.8	3.7	1411	4	US-09-496-692-5	Sequence 5, Appli
	42	76.8	3.7	1411	4	US-10-000-273-5	Sequence 5, Appli
	43	76.6	3.7	578	4	US-09-602-877A-95	Sequence 95, Appl
	44	76.6	3.7	949	4	US-09-489-847-35	Sequence 35, Appl
	45	76.6	3.7	1683	3	US-09-347-803-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-509-902A-3
; Sequence 3, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1085
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-3

Query Match 48.8%; Score 1004.4; DB 4; Length 1085;
Best Local Similarity 99.9%; Pred. No. 4.8e-207;
Matches 1005; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 139

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 199

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 259

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 319

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 379

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 439

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 499

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 559

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 619

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 679

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 739

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 799

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 859

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 919

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 979

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1039

Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1006
                ||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1085


RESULT 2
US-09-220-132-135/c
; Sequence 135, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  METHODS AND COMPOSITONS FOR THE IDENTIFICATION AND ASSESSMENT
;  TITLE OF INVENTION:  OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCE
;  FILE REFERENCE:  07334-074001
;  CURRENT APPLICATION NUMBER: US/09/220,132
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  US 60/079,303
;  PRIOR FILING DATE:  1998-03-25
;  PRIOR APPLICATION NUMBER:  US 60/068,821
;  PRIOR FILING DATE:  1997-12-24
;  NUMBER OF SEQ ID NOS:  191
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 135
;   LENGTH: 2559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-220-132-135

  Query Match             24.5%;  Score 505.2;  DB 4;  Length 2559;
  Best Local Similarity   94.6%;  Pred. No. 1.4e-99;
  Matches  566;  Conservative    0;  Mismatches   28;  Indels    4;  Gaps    4;

Qy        604 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 663
                || | |||    ||| |     ||   ||||||||||||||||||||||||||||||||

Db       1634 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 1575

Qy        664 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 723
                |||||||||||||||||||||||||||||||||||| ||||||||||||  ||||||||||
Db       1574 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT-GGACAAGCACGCGTCGCCAGCCTAC 1516

Qy        724 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 783
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 1456

Qy        784 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 843
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 1396

Qy        844 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 903
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db       1395 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGCGCTCTCG 1336

Qy        904 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 963
                || |  | ||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1335 CCCTCGCCGCGCTGTCTGGTTCCGTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 1276

Qy        964 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 1023
                | |||||  |||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1275 G-CACAGCGATCCTCCTGCACCCCTGGCTGCGACA-GACCCGATGCCCTTAGCCCCAACC 1218

Qy       1024 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1083
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1217 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA-CC 1159

Qy       1084 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1143
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1158 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1099

Qy       1144 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGA 1201
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1098 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGA 1041


RESULT 3
US-08-146-421-4/c
; Sequence 4, Application US/08146421
; Patent No. 5543499
;  GENERAL INFORMATION:
;    APPLICANT:  BREWER, GARY
;    TITLE OF INVENTION:  DNA SEQUENCE ENCODING A POLYPEPTIDE WITH
;    TITLE OF INVENTION:  ANTI-TUMOR PROPERTIES
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DILWORTH & BARRESE
;      STREET:  4350 LA JOLLA VILLAGE DRIVE, SUITE 300
;      CITY:  SAN DIEGO
;      STATE:  CALIFORNIA
;      COUNTRY:  U.S.A.
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/146,421
;      FILING DATE:  29-OCT-1993
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PEPPER PH.D., FREDERICK W.
;      REGISTRATION NUMBER:  31,286
;      REFERENCE/DOCKET NUMBER:  489-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-4410
;      TELEFAX:  619-453-2839
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  2562 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  246..1106
US-08-146-421-4

  Query Match             24.5%;  Score 505.2;  DB 1;  Length 2562;
  Best Local Similarity   94.6%;  Pred. No. 1.4e-99;
  Matches  566;  Conservative    0;  Mismatches   28;  Indels    4;  Gaps    4;

Qy        604 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 663
              || | |||     ||| |       ||   ||||||||||||||||||||||||||||||
Db       1637 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 1578

Qy        664 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGACAAGCACGCGTGCCCAGCCTAC 723
              |||||||||||||||||||||||||||||||| ||||||||||||||  ||||||||||
Db       1577 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT-GGACAAGCACGCGTCGCCAGCCTAC 1519

Qy        724 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1518 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 1459

Qy        784 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1458 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 1399

Qy        844 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 903
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db       1398 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGCGCTCTCG 1339

Qy        904 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 963
                || |   |||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1338 CCCTCGCCGCGCTGTCTGGTTCCGTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 1279

Qy        964 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 1023
              | |||||  |||||||||||||||||||||||||| |||||||||||||||||||||||||
Db       1278 G-CACAGCGATCCTCCTGCACCCCTGGCTGCGACA-GACCCGATGCCCTTAGCCCCAACC 1221

Qy       1024 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1220 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA-CC 1162

Qy       1084 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1161 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1102

Qy       1144 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGA 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGA 1044

RESULT 4
US-08-913-050A-2
; Sequence 2, Application US/08913050A
; Patent No. 5827726
;  GENERAL INFORMATION:
;    APPLICANT:  NEZU, Jun-ichi
;    TITLE OF INVENTION:  DNA ENCODING PROTEIN KINASE
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK, P.L.L.C.
;      STREET:  419 7th Street N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/913,050A
;      FILING DATE:  05-SEP-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 57104/1995
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP PCT/JP96/00660
;      FILING DATE:  15-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YUN, Allen C.
;      REGISTRATION NUMBER:  37,971
;      REFERENCE/DOCKET NUMBER:  NEZU=4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1302 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-913-050A-2

  Query Match              4.1%;  Score 85;  DB 1;  Length 1302;
  Best Local Similarity   48.1%;  Pred. No. 2.6e-09;
  Matches  241;  Conservative    0;  Mismatches  260;  Indels    0;  Gaps    0;

Qy        503 ACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGC 562
              | ||| ||||  | |      ||||   || |   ||||  | ||| ||      || ||||| |
Db        437 AGCGTTTCCCAGTGTGCCAGGCCCACGGGTACTTCTGTCAGCTGATTGACGGCCTGGAGT 496

Qy        563 ACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTG 622
               ||       ||  ||   | ||        ||  ||||||  |      |   ||  |
Db        497 ACCTGCATAGCCAGGGCATTGTGCACAAGGACATCAAGCCGGGGAACCTGCTGCTCACCA 556

Qy        623 ACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 682
                | |||  |      || |            | | |||  ||  |          | |   | |
Db        557 CCGGTGGCACCCTCAAAATCTCCGACCTGGGCGTGGCCGAGGCACTGCACCCGTTCGCGG 616

Qy        683 CAGATGATTCCCTGTGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCA 742
               | || ||  ||     ||  ||  ||  |    | ||| ||  |  |    |  || |||||    ||
Db        617 CGGACGACACCTGCCGGACCAGCCAGGGCTCCCCGGCTTTCCAGCCGCCCGAGATTGCCA 676

Qy        743 GCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCT 802
                |      | || |  |  |  ||| |||            |  ||    |||||          || ||   |  ||||
Db        677 ACGGCCTGGACACCTTCTCCGGCTTCAAGGTGGACATCTGGTCGGCTGGGGTCACCCTCT 736

Qy        803 TCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCA 862
               || |||          |  || |   ||||||||| |        |            |   |  |  ||   |
Db        737 ACAACATCACCACGGGTCTGTACCCCTTCGAAGGGGACAACATCTACAAGTTGTTTGAGA 796

Qy        863 AGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGG 922
              | |||  |       |||   ||||||| |  |  |        ||||      |       |||
Db        797 ACATCGGGAAGGGGAGCTACGCCATCCCGGGCGACTGTGGCCCCCCGCTCTCTGACCTGC 856

Qy        923 TTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGC 982
              |      |  | |||        || || ||  |  || || |  |          ||||  | ||
Db        857 TGAAAGGGATGCTTGAGTACGAACCGGCCAAGAGGTTCTCCATCCGGCAGATCCGGCAGC 916

Qy        983 ACCCCTGGCTGCGACAGGACC 1003
               || |||| | ||  || | |
Db        917 ACAGCTGGTTCCGGAAGAAAC 937

RESULT 5
US-09-016-434-1146
; Sequence 1146, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1146:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1302 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g1480860
US-09-016-434-1146

  Query Match             4.1%;  Score 85;  DB 4;  Length 1302;
  Best Local Similarity   48.1%;  Pred. No. 2.6e-09;
  Matches  241;  Conservative    0;  Mismatches  260;  Indels    0;  Gaps    0;

Qy        503 ACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGC 562
              | ||| |||| | |    ||||   || |  |||| | ||| ||     || ||||| |
Db        437 AGCGTTTCCCAGTGTGCCAGGCCCACGGGTACTTCTGTCAGCTGATTGACGGCCTGGAGT 496

Qy        563 ACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTG 622
              ||   ||    || ||  | || |    ||  ||||| | |   | |  | || |
Db        497 ACCTGCATAGCCAGGGCATTGTGCACAAGGACATCAAGCCGGGGAACCTGCTGCTCACCA 556

Qy        623 ACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 682
              | |||  |     ||  |        | | |||  ||  |       | |   | |
Db        557 CCGGTGGCACCCTCAAAATCTCCGACCTGGGCGTGGCCGAGGCACTGCACCCGTTCGCGG 616

Qy        683 CAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCA 742
              | || ||  ||    ||  ||  || |  | ||| || | |   |   || |||||   ||
Db        617 CGGACGACACCTGCCGGACCAGCCAGGGCTCCCCGGCTTTCCAGCCGCCCGAGATTGCCA 676

Qy        743 GCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCT 802
              |   | || | | | ||| |||      |  ||  |||||      || ||  | ||||
Db        677 ACGGCCTGGACACCTTCTCCGGCTTCAAGGTGGACATCTGGTCGGCTGGGGTCACCCTCT 736

Qy        803 TCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCA 862
              || |||     | || |  |||||||| | |     |    |  | | || |  |
Db        737 ACAACATCACCACGGGTCTGTACCCCTTCGAAGGGGACAACATCTACAAGTTGTTTGAGA 796

Qy        863 AGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGG 922
              | ||| |    |||  ||||||| | || |  | |        ||||   |      |||
Db        797 ACATCGGGAAGGGGAGCTACGCCATCCCGGGCGACTGTGGCCCCCCGCTCTCTGACCTGC 856

Qy        923 TTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGC 982
              |      |  | |||      || || ||  |  || || |  |      |||| | |||
Db        857 TGAAAGGGATGCTTGAGTACGAACCGGCCAAGAGGTTCTCCATCCGGCAGATCCGGCAGC 916

Qy        983 ACCCCTGGCTGCGACAGGACC 1003
              ||  |||| | ||  || | |
Db        917 ACAGCTGGTTCCGGAAGAAAC 937


RESULT 6
US-09-559-023-1
; Sequence 1, Application US/09559023
; Patent No. 6551796
; GENERAL INFORMATION:
;  APPLICANT: Abramson, Ruth
;  APPLICANT:  Leal-Pinto, Edgar
;  APPLICANT:  Lipkowitz, Michael
;  TITLE OF INVENTION: NUCLEIC ACID ENCODING URATE TRANSPORTER
;  TITLE OF INVENTION:  AND METHODS OF USE THEREOF
;  FILE REFERENCE: 070165.0574
;  CURRENT APPLICATION NUMBER: US/09/559,023
;  CURRENT FILING DATE:  2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/221,898
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: US 60/099,752
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: US 60/070,215
;  PRIOR FILING DATE: 1997-12-31
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1545
;   TYPE: DNA
;   ORGANISM: Rat
US-09-559-023-1

  Query Match             4.1%;  Score 84.6;  DB 4;  Length 1545;
  Best Local Similarity   68.4%;  Pred. No. 3.4e-09;
  Matches  117;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Qy       1889 CTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACTCTAGGTTTTGGATACCATG 1948
              || | |  ||| |      |||||| | ||  |||  | |||        |      |
Db       1375 CTCCAGCTGACCCAGAATGTCCACATTACCTTTTCCCCAATCTTTCCCAATGCCCATAAA 1434

Qy       1949 AGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAA 2008
              |  | | ||  |  ||    || |||  |||  | ||  |  ||| ||||||||||||||||
Db       1435 ATAAAGAATATCAACGCTTGTCTACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1494

Qy       2009 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1545


RESULT 7
US-09-996-243-300
; Sequence 300, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi,Avi J.
;  APPLICANT:  Baker,Kevin P.
;  APPLICANT:  Botstein,David
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Eaton,Dan L.
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Fong,Sherman
;  APPLICANT:  Gerber,Hanspeter
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,J.Christopher
;  APPLICANT:  Gurney,Austin L.
```

; APPLICANT: Kljavin,Ivar J.
; APPLICANT: Napier,Mary A.
; APPLICANT: Pan,James
; APPLICANT: Paoni,Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 4.0%; Score 82.8; DB 4; Length 1674;
Best Local Similarity 58.2%; Pred. No. 8.5e-09;
Matches 164; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

```
Qy       1779 TGAGGGTCCAGGCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAG-GGACA 1837
             ||  | | ||||   || |      |||||  ||  |  ||||  |  ||  | |  |||
Db       1368 TGGAGCTGCAGGATGGTGAGGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAACA 1427

Qy       1838 GGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGG 1897
             |  |   | |||  |||   || ||||  |     | | |     |||||  |  || | |
Db       1428 GCTTGCTGAAGGAGCTGAGATGAGGCCCCACACCAGGACTGGACTGGGATGCCGCTAGTG 1487

Qy       1898 ACACTTGGGGTCCACAATCCCAGGTCCATACTCTAGGTTTTGGATACCATGAGTATGTAT 1957
             |  ||  |||    ||       |      | |||  |      |||  | | ||    ||
Db       1488 AGGCTGAGGGGTGCCAGCGTGGGTGGGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTG 1547

Qy       1958 GTTTACCTGTGCCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
             | ||  |  |       |||   |||      ||| ||||||||||||||||||||||||
Db       1548 GCTTGGCCATTAAATGGAAACCTGAAGGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1607

Qy       2018 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2059
              |||||||||||||||||||||||||||||||||||||||||
Db       1608 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1649
```

RESULT 8
US-09-996-243-302
; Sequence 302, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi,Avi J.
; APPLICANT: Baker,Kevin P.
; APPLICANT: Botstein,David
; APPLICANT: Desnoyers,Luc
; APPLICANT: Eaton,Dan L.
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Fong,Sherman
; APPLICANT: Gerber,Hanspeter
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,J.Christopher
; APPLICANT: Gurney,Austin L.
; APPLICANT: Kljavin,Ivar J.
; APPLICANT: Napier,Mary A.
; APPLICANT: Pan,James
; APPLICANT: Paoni,Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05

```
;  PRIOR APPLICATION NUMBER: 60/088212
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088217
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088655
;  PRIOR FILING DATE: 1998-06-09
;  PRIOR APPLICATION NUMBER: 60/088734
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088738
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088742
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088810
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088824
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088826
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088858
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/088861
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/088876
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/089105
;  PRIOR FILING DATE: 1998-06-12
;  PRIOR APPLICATION NUMBER: 60/089440
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089512
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089514
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089532
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089538
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089598
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089599
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089600
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089653
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/089907
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/089908
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/089947
;  PRIOR FILING DATE: 1998-06-19
;  PRIOR APPLICATION NUMBER: 60/089948
;  PRIOR FILING DATE: 1998-06-19
;  PRIOR APPLICATION NUMBER: 60/089952
;  PRIOR FILING DATE: 1998-06-19
;  PRIOR APPLICATION NUMBER: 60/090246
;  PRIOR FILING DATE: 1998-06-22
;  PRIOR APPLICATION NUMBER: 60/090252
;  PRIOR FILING DATE: 1998-06-22
;  PRIOR APPLICATION NUMBER: 60/090254
;  PRIOR FILING DATE: 1998-06-22
;  PRIOR APPLICATION NUMBER: 60/090349
;  PRIOR FILING DATE: 1998-06-23
;  PRIOR APPLICATION NUMBER: 60/090355
;  PRIOR FILING DATE: 1998-06-23
;  PRIOR APPLICATION NUMBER: 60/090429
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090431
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090435
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090444
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090445
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090472
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090535
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090540
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090542
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090676
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090678
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090690
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090694
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090695
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090696
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090862
;  PRIOR FILING DATE: 1998-06-26
;  PRIOR APPLICATION NUMBER: 60/090863
;  PRIOR FILING DATE: 1998-06-26
;  PRIOR APPLICATION NUMBER: 60/091360
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091478
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091544
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091519
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091626
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091633
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091978
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/091982
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/092182
;  PRIOR FILING DATE: 1998-07-09

  Query Match             4.0%;  Score 82.4;  DB 4;  Length 2136;
  Best Local Similarity   85.2%;  Pred. No. 1.1e-08;
  Matches   92;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1952 ATGTATGTTTACCTGTGCCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAAA 2011
              | ||   | |   | || ||||||||  ||||||   ||| |||||||||||||||||||
Db       2027 AGGTGATTGTGATTTTGACTAATAAAAAAGAATTTGTAAAAAAAAAAAAAAAAAAAAAAA 2086

Qy       2012 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2059
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2087 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2134

RESULT 9
US-08-749-902-2
; Sequence 2, Application US/08749902
; Patent No. 5985635
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    TITLE OF INVENTION:  NOVEL HUMAN SERINE/THREONINE
;    TITLE OF INVENTION:  PROTEIN KINASES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
```

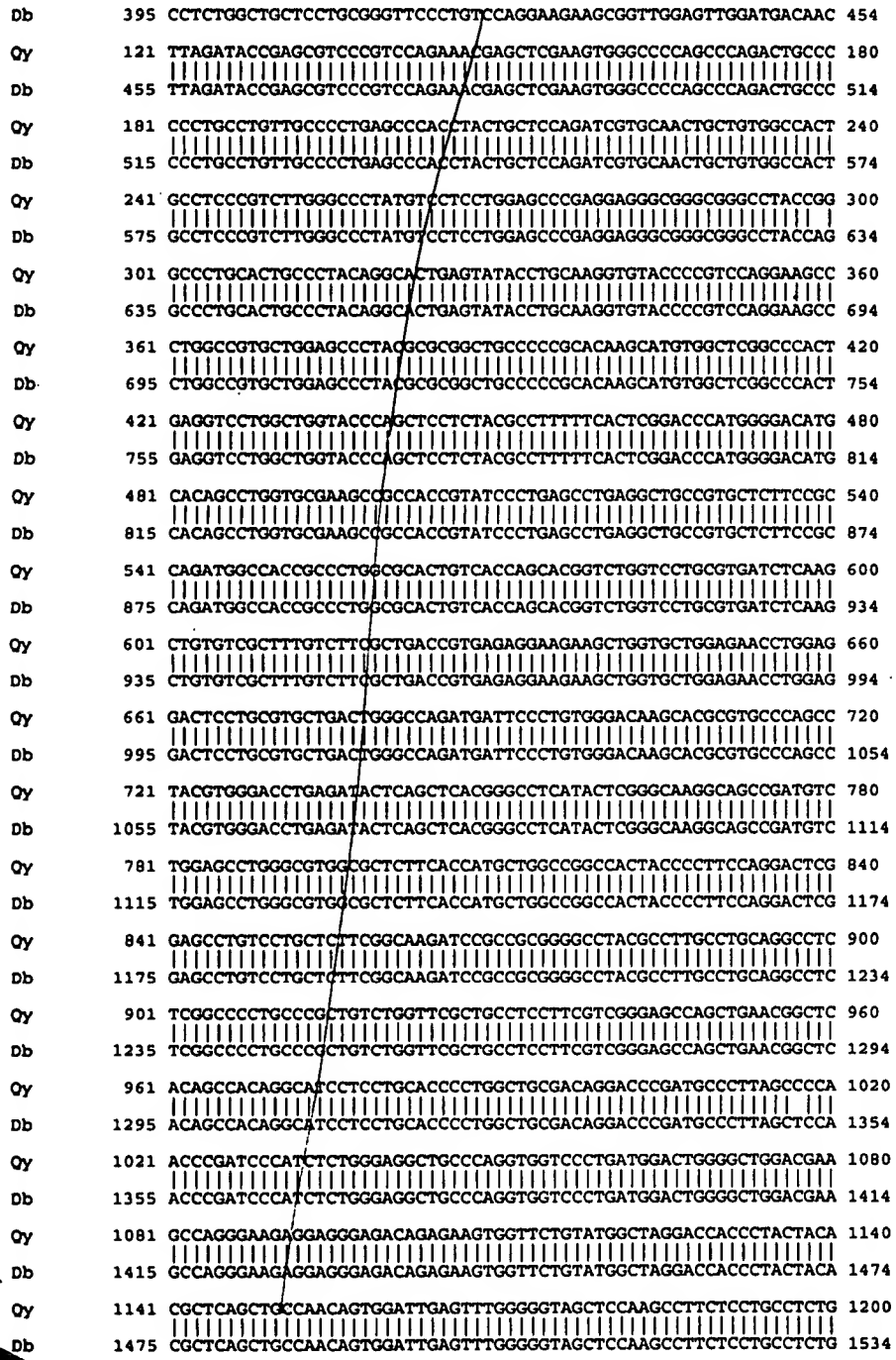
```
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/749,902
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0150 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1466 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  Consensus
US-08-749-902-2

  Query Match             4.0%;  Score 81.6;  DB 2;  Length 1466;
  Best Local Similarity   47.3%;  Pred. No. 1.5e-08;
  Matches  237;  Conservative    0;  Mismatches  264;  Indels    0;  Gaps    0;

Qy        503 ACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGC 562
              | ||| |||| | |    ||||   || |  |||| | ||| ||     || | ||| |
Db        536 AGCGTTTCCCAGTGTGCCAGGCCCACGGGTACTTCTGTCAGCTGATTGACGGCNTGGGGT 595

Qy        563 ACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTG 622
              ||   ||    || ||  | || |      |   |||||| | |   | | |  || |
Db        596 ACCTGCATNGCCAGGGNATTGTGCACAAGGGCATCAAGCCGGGGAACCTGCTGCTCACCA 655

Qy        623 ACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 682
              | |||  |     ||  |          |  | |||  ||  |       | |   |
Db        656 CCGGTGGCACCCTCAAAATCTCCGACCTGGGCGTGGCCGAGGCACTGCACCCGTTCGCNG 715

Qy        683 CAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCA 742
              | || ||  ||    ||  ||  || |  | ||| || | |   |   || || ||   ||
Db        716 CGGACGACACCTGCCGGACCAGCCAGGGCTCCCCGGCTTTCCAGCCGCCCGANATTGCCA 775

Qy        743 GCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCT 802
              |   | || | | | ||| |||       |  ||  |||||       || ||  | ||||
Db        776 ACGGCCTGGACACCTTCTCCGGCTTCAAGGTGGACATCTGGTCGGCTGGGGTCACCCTCT 835

Qy        803 TCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCA 862
              || |||     | || |  |||||||||| | |       |       | | | || | |
Db        836 ACAACATCACCACGGGTCTGTACCCCTTCGAAGGGGACAACATCTACAAGTTGTTTGAGA 895

Qy        863 AGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGG 922
              | ||| |    |||  ||||||| | || |  | |       ||||   |      |||
Db        896 ACATCGGGAAGGGGAGCTACGCCATCCCGGGCGACTGTGGCCCCCCGCTCTCTGACCTGC 955

Qy        923 TTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGC 982
              |      | | |||      || || ||  |  || || |   |       ||||  | ||
Db        956 TGAAAGGGATGCTTGAGTACGAACCGGCCAAGAGGTTCTCCATCCGGCAGATCCGGCAGC 1015

Qy        983 ACCCCTGGCTGCGACAGGACC 1003
              ||  |||| | ||  | | | |
Db       1016 ACAGCTGGTTCCGGAAGAAAC 1036
```

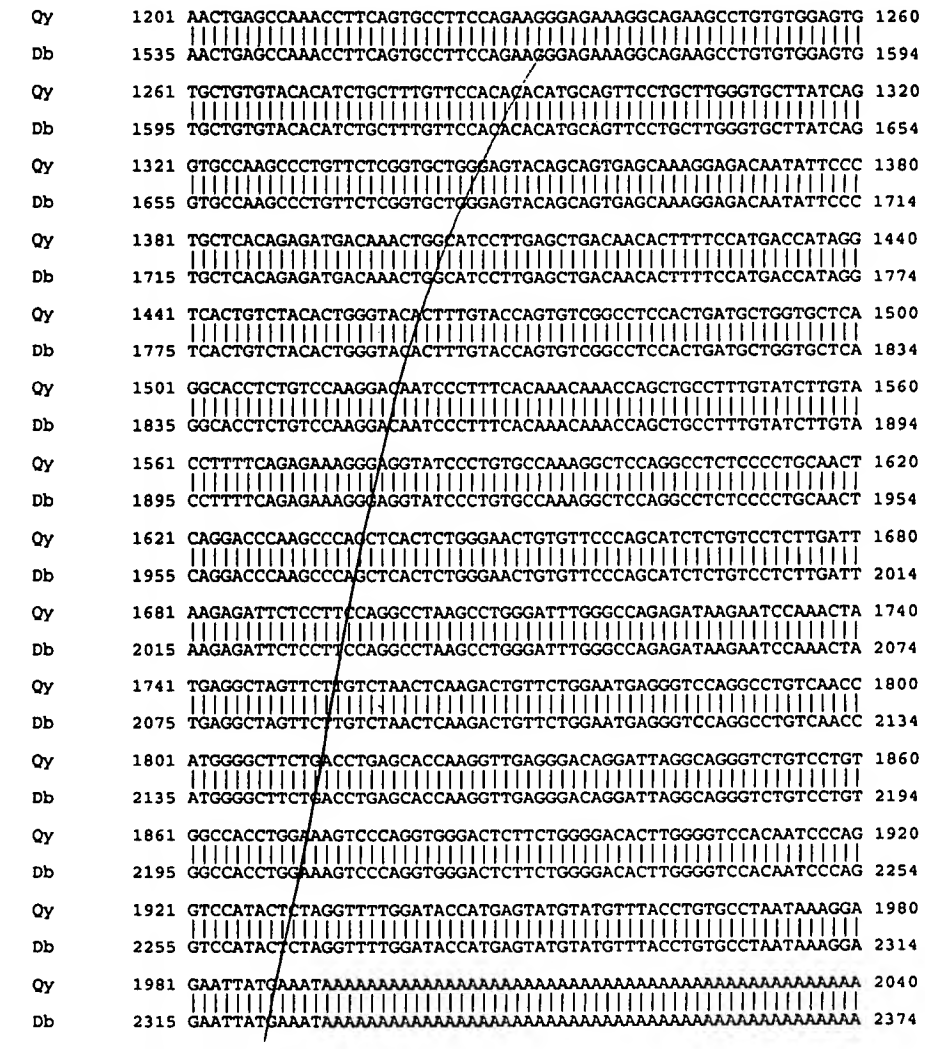
```
RESULT 10
US-09-489-847-38
; Sequence 38, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al
;  TITLE OF INVENTION: 98 Human Secreted Proteins
;  FILE REFERENCE: PZ031P1
;  CURRENT APPLICATION NUMBER: US/09/489,847
;  CURRENT FILING DATE: 2000-01-24
;  EARLIER APPLICATION NUMBER: PCT/US99/17130
;  EARLIER FILING DATE: 1999-07-29
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  EARLIER APPLICATION NUMBER: 60/095,486
;  EARLIER FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/096,319
;  EARLIER FILING DATE: 1998-08-12
;  EARLIER APPLICATION NUMBER: 60/095,454
;  EARLIER FILING DATE: 1998-08-06
;  EARLIER APPLICATION NUMBER: 60/095,455
;  EARLIER FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 376
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
;   LENGTH: 1048
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-489-847-38

  Query Match             3.9%;  Score 79.8;  DB 4;  Length 1048;
  Best Local Similarity   87.9%;  Pred. No. 3.2e-08;
  Matches   87;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1961 TACCTGTGCCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAAA 2020
              ||  | | ||| || ||||| ||  |  ||| ||||||||||||||||||||||||||||
Db        911 TAAATATTCCTTATTAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 970

Qy       2021 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2059
              ||||||||||||||||||||||||||||||||||||||||
Db        971 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1009


RESULT 11
US-09-996-243-245
; Sequence 245, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi,Avi J.
;  APPLICANT:  Baker,Kevin P.
;  APPLICANT:  Botstein,David
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Eaton,Dan L.
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Fong,Sherman
;  APPLICANT:  Gerber,Hanspeter
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,J.Christopher
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Kljavin,Ivar J.
;  APPLICANT:  Napier,Mary A.
;  APPLICANT:  Pan,James
;  APPLICANT:  Paoni,Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Watanabe, Colin K.
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
```

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25

```
;  PRIOR APPLICATION NUMBER: 60/090694
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090695
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090696
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090862
;  PRIOR FILING DATE: 1998-06-26
;  PRIOR APPLICATION NUMBER: 60/090863
;  PRIOR FILING DATE: 1998-06-26
;  PRIOR APPLICATION NUMBER: 60/091360
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091478
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091544
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091519
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091626
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091633
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091978
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/091982
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/092182
;  PRIOR FILING DATE: 1998-07-09

  Query Match             3.9%;  Score 79.6;  DB 4;  Length 485;
  Best Local Similarity   72.5%;  Pred. No. 2.7e-08;
  Matches  103;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       1918 CAGGTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAA 1977
              | |  ||| |   ||||| |||||     |   |   ||   |||   |       || |||
Db        343 CTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTGGTTCTTTTGTTTTCCAAAAAAAAA 402

Qy       1978 GGAGAATTATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2037
              | ||  |  ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 462

Qy       2038 AAAAAAAAAAAAAAAAAAAA 2059
              ||||||||||||||||||||
Db        463 AAAAAAAAAAAAAAAAAAAA 484


RESULT 12
US-08-157-101A-4
; Sequence 4, Application US/08157101A
; Patent No. 5808032
;  GENERAL INFORMATION:
;    APPLICANT:  KURIHARA, TATSUYA
;    APPLICANT:  MATSUKURA, SHIGEKAZU
;    APPLICANT:  TSURUOKA, NOBUO
;    APPLICANT:  ARIMA, KENJI
;    APPLICANT:  NISHIHARA, TATSURO
;    TITLE OF INVENTION:  ANTI-HBs ANTIBODY GENES AND EXPRESSION
;    TITLE OF INVENTION:  PLASMIDS THEREFOR
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PILLSBURY, MADISON & SUTRO
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/157,101A
;      FILING DATE:  05-APR-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TITUS, MARLANA K
;      REGISTRATION NUMBER:  35843
;      REFERENCE/DOCKET NUMBER:  9437/204199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861-3711
;      TELEFAX:  202-822-0944
;      TELEX:  6714627 CUCH
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1066 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-157-101A-4

  Query Match             3.9%;  Score 79.6;  DB 1;  Length 1066;
  Best Local Similarity   74.6%;  Pred. No. 3.6e-08;
  Matches  100;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       1926 TACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATT 1985
              ||  |||   || |   |  |||| ||  ||   | |     |  | || |||  | ||
Db        893 TATGCTAATGTTGGAGGAGAATGAATAAATAAAGTGAATCTTTGCAAAAAAAAAAAAAAA 952

Qy       1986 ATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2045
              | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1012

Qy       2046 AAAAAAAAAAAAAA 2059
              ||||||||||||||
Db       1013 AAAAAAAAAAAAAA 1026


RESULT 13
US-09-336-536-66
; Sequence 66, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
;   LENGTH: 1927
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: all "n" positions
;   OTHER INFORMATION: n=a, c, g, or t
US-09-336-536-66

  Query Match             3.9%;  Score 79.6;  DB 4;  Length 1927;
  Best Local Similarity   89.5%;  Pred. No. 4.3e-08;
  Matches   85;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1965 TGTGCCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2024
              | ||| | | |||||| |||     ||| ||||||||||||||||||||||||||||||||
Db       1766 TTTGCATTAAAAAGGAAAATCCANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1825

Qy       2025 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2059
              |||||||||||||||||||||||||||||||||||
Db       1826 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1860
```

```
RESULT 14
US-09-328-965-1
; Sequence 1, Application US/09328965
; Patent No. 6501008
; GENERAL INFORMATION:
;  APPLICANT: Nevins, Donald J.
;  APPLICANT: Simmons, Carl
;  APPLICANT: The Regents of the University of California
;  TITLE OF INVENTION: Endo- and Exo-Glucanases and Gene
;  FILE REFERENCE: 02307O-096600US
;  CURRENT APPLICATION NUMBER: US/09/328,965
;  CURRENT FILING DATE: 1999-06-09
;  EARLIER APPLICATION NUMBER: US 60/088,780
;  EARLIER FILING DATE: 1998-06-10
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1091
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: maize coleoptile endo-1,3;1,4-beta glucanase cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (68)..(979)
;   OTHER INFORMATION: endo-1,3;1,4-beta glucanase
US-09-328-965-1

  Query Match             3.8%;  Score 79.2;  DB 4;  Length 1091;
  Best Local Similarity   75.0%;  Pred. No. 4.4e-08;
  Matches   99;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       1928 CTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTAT 1987
              ||||| || ||   |  | |   | | |  |||   ||  |  |||||||   | |
Db        940 CTCTATGTATTTCTTTTCTTTTCTTTTTGCTTTTTTATGATCGCAATAAAGTTCAGTAGG 999

Qy       1988 GAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2047
              |  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1059

Qy       2048 AAAAAAAAAAAA 2059
              ||||||||||||
Db       1060 AAAAAAAAAAAA 1071


RESULT 15
US-09-370-253-1
; Sequence 1, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen M.
;  APPLICANT: Sakai, Hajime
;  APPLICANT: Thorpe, Catherine J.
;  TITLE OF INVENTION: Amino Acid Transporters
;  FILE REFERENCE: BB-1200
;  CURRENT APPLICATION NUMBER: US/09/370,253
;  CURRENT FILING DATE: 1999-08-09
;  EARLIER APPLICATION NUMBER: 60/097,222
;  EARLIER FILING DATE: August 20, 1998
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
;   LENGTH: 1882
;   TYPE: DNA
;   ORGANISM: Hordeum vulgare
US-09-370-253-1

  Query Match             3.8%;  Score 79.2;  DB 3;  Length 1882;
  Best Local Similarity   83.3%;  Pred. No. 5.2e-08;
  Matches   90;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1952 ATGTATGTTTACCTGTGCCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAA 2011
              |||||||  | | | | | |   |||  | ||  |  ||| ||||||||||||||||||

Db       1760 ATGTATGGGTTCTTTTTCTTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1819

Qy       2012 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1867


Search completed: January 16, 2004, 16:49:41
Job time : 135.44 secs
```



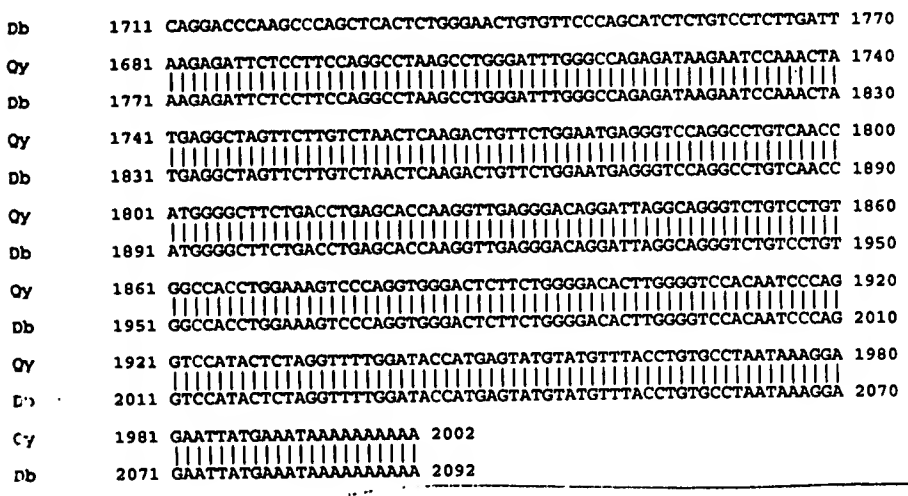


RESULT 2
US-10-098-841-271
; Sequence 271, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing

; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 271
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (139)..(1215)
US-10-098-841-271

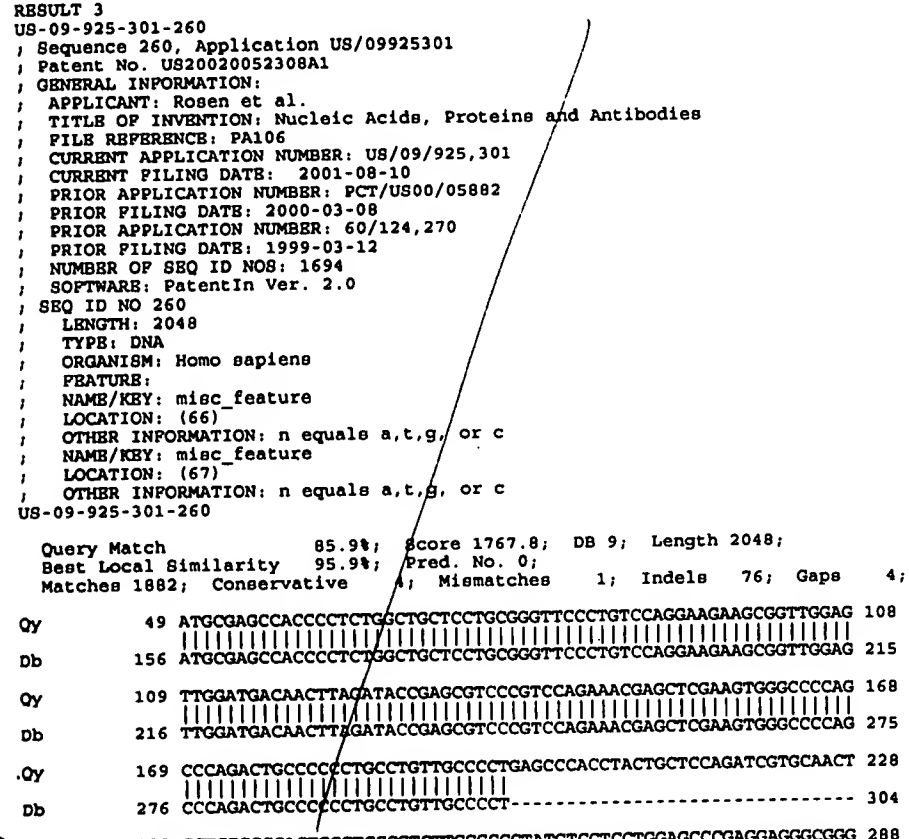
Query Match 97.0%; Score 1997.2; DB 14; Length 2092;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

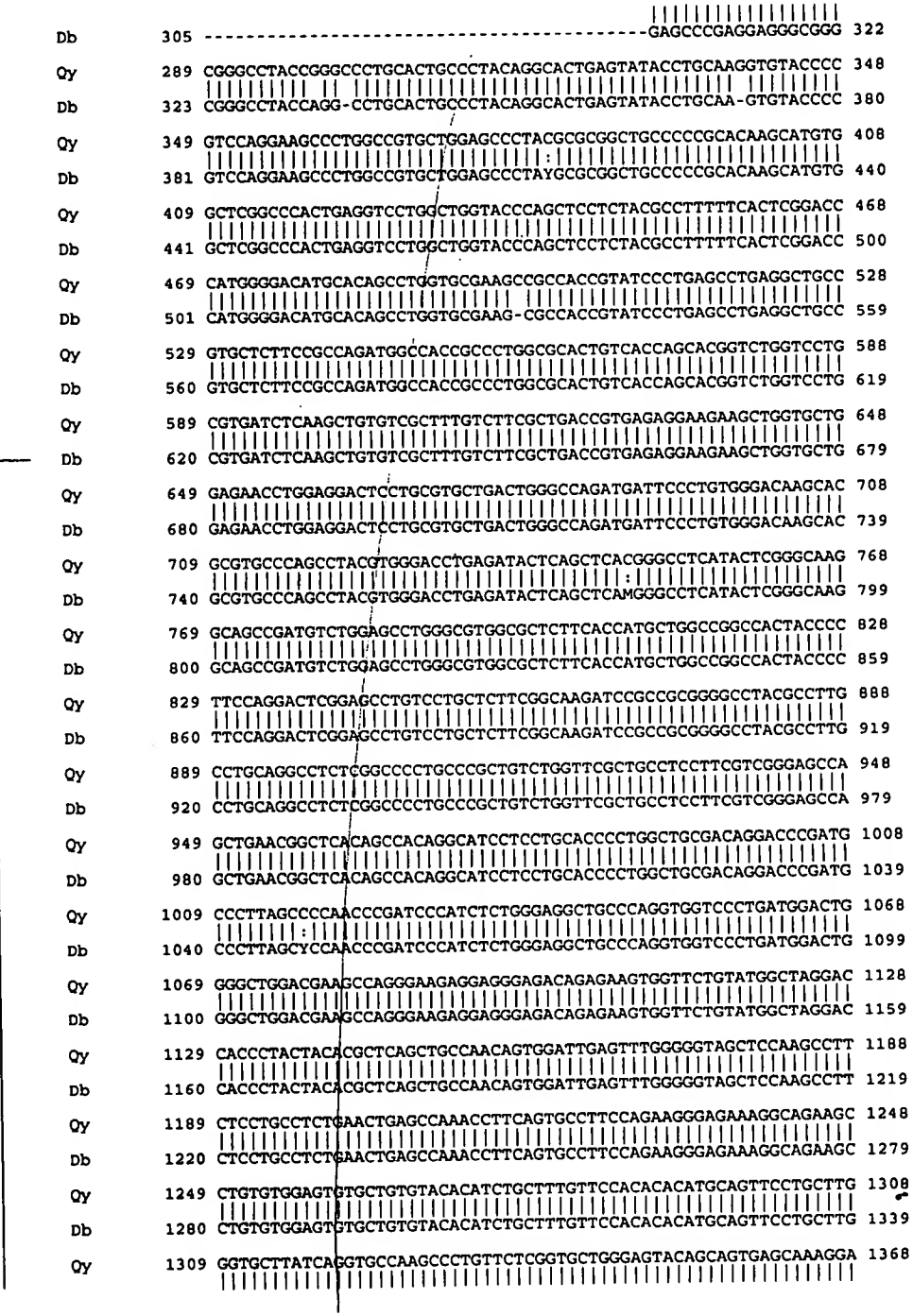
```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 150

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 210

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 270

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 330

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        331 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 390

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        391 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAGGGTGTACCCCGTCCAGGAAGCC 450

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        451 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 510

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 570

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 630

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 690

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 750

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 810

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        811 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 870

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        871 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 930

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        931 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 990

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        991 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1050

Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1051 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1110

Qy       1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1111 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1170

Qy       1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1171 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1230

Qy       1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1231 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1290

Qy       1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1291 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1350

Qy       1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1351 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1410

Qy       1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1411 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1470

Qy       1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1471 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1530

Qy       1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1531 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1590

Qy       1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1591 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1650

Qy       1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1710

Qy       1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



RESULT 3
US-09-925-301-260
; Sequence 260, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 260
; LENGTH: 2048
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (66)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (67)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-260

Query Match 85.9%; Score 1767.8; DB 9; Length 2048;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1882; Conservative 4; Mismatches 1; Indels 76; Gaps 4;





GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 14:49:15 ; Search time 703.201 Seconds
(without alignments)
10320.639 Million cell updates/sec

Title: US-09-909-474D-1
Perfect score: 2059
Sequence: 1 gctctgagccccggcggcgc.....aaaaaaaaaaaaaaaaaaaa 2059

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2051.8	99.7	2389	9	US-09-799-875-7	Sequence 7, Appli
2	1997.2	97.0	2092	14	US-10-098-841-271	Sequence 271, App
3	1767.8	85.9	2048	9	US-09-925-301-260	Sequence 260, App
4	1070.8	52.0	1074	9	US-09-799-875-9	Sequence 9, Appli
5	1004.4	48.8	1085	15	US-10-024-828-3	Sequence 3, Appli
6	533	25.9	541	10	US-09-919-580-580	Sequence 580, App
c 7	426	20.7	426	10	US-09-867-701-10282	Sequence 10282, A
8	330.2	16.0	396	11	US-09-803-719-519	Sequence 519, App
c 9	315.4	15.3	353	13	US-09-873-367C-341	Sequence 341, App
10	268.8	13.1	562	9	US-09-925-301-525	Sequence 525, App
11	263.2	12.8	1909	12	US-10-291-172-116	Sequence 116, App
12	263.2	12.8	1909	15	US-10-119-926-73	Sequence 73, Appl
13	253	12.3	4336	15	US-10-084-817-300	Sequence 300, App
14	251.4	12.2	3280	14	US-10-044-090-269	Sequence 269, App
15	247	12.0	3324	13	US-10-240-965-168	Sequence 168, App
16	232.4	11.3	818	15	US-10-228-263-1	Sequence 1, Appli
17	222.2	10.8	621	15	US-10-228-263-3	Sequence 3, Appli
18	207.6	10.1	1943	12	US-10-291-172-492	Sequence 492, App
c 19	165	8.0	498	10	US-09-917-800A-991	Sequence 991, App
c 20	158	7.7	562	9	US-09-864-761-12987	Sequence 12987, A
c 21	156	7.6	417	9	US-09-864-761-29550	Sequence 29550, A
22	138.8	6.7	4463	12	US-10-062-674-1704	Sequence 1704, Ap
23	88.2	4.3	2288	13	US-10-266-829-19	Sequence 19, Appl
24	88.2	4.3	2288	15	US-10-106-698-464	Sequence 464, App
25	85.2	4.1	2260	13	US-10-220-891-11	Sequence 11, Appl
26	84.6	4.1	1545	10	US-09-917-800A-1591	Sequence 1591, Ap
27	84	4.1	2511	13	US-10-136-547-13	Sequence 13, Appl
28	84	4.1	2511	15	US-10-091-752A-13	Sequence 13, Appl
29	83.6	4.1	993	12	US-10-443-201-53	Sequence 53, Appl
30	83.4	4.1	2120	11	US-09-746-783-163	Sequence 163, App
c 31	83.2	4.0	239	10	US-09-960-352-11438	Sequence 11438, A
32	83.2	4.0	929	12	US-10-264-049-433	Sequence 433, App
33	83	4.0	326	10	US-09-924-401-51	Sequence 51, Appl
34	82.8	4.0	1674	9	US-09-989-722-300	Sequence 300, App
35	82.8	4.0	1674	9	US-09-989-723-300	Sequence 300, App
36	82.8	4.0	1674	9	US-09-989-279-300	Sequence 300, App
37	82.8	4.0	1674	9	US-09-989-727-300	Sequence 300, App
38	82.8	4.0	1674	10	US-09-989-731-300	Sequence 300, App
39	82.8	4.0	1674	10	US-09-989-732-300	Sequence 300, App
40	82.8	4.0	1674	10	US-09-991-073-300	Sequence 300, App
41	82.8	4.0	1674	10	US-09-990-442-300	Sequence 300, App
42	82.8	4.0	1674	10	US-09-991-163-300	Sequence 300, App
43	82.8	4.0	1674	10	US-09-993-604-300	Sequence 300, App
44	82.8	4.0	1674	10	US-09-990-456-300	Sequence 300, App
45	82.8	4.0	1674	10	US-09-989-721-300	Sequence 300, App

ALIGNMENTS

RESULT 1
US-09-799-875-7
; Sequence 7, Application US/09799875
; Patent No. US20020034780A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. US20020034780A1el Human Protein Kinases and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2389
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (383)...(1456)
US-09-799-875-7

Query Match 99.7%; Score 2051.8; DB 9; Length 2389;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2053; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 394

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        395 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 454
Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
Db        455 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 514
Qy        181 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
Db        515 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 574
Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
Db        575 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 634
Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
Db        635 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 694
Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
Db        695 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 754
Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
Db        755 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 814
Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
Db        815 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 874
Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
Db        875 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 934
Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
Db        935 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 994
Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
Db        995 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 1054
Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
Db       1055 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 1114
Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
Db       1115 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 1174
Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
Db       1175 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 1234
Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
Db       1235 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1294
Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
Db       1295 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCTCCA 1354
Qy       1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
Db       1355 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1414
Qy       1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
Db       1415 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1474
Qy       1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
Db       1475 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1534
```

```
Qy       1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
Db       1535 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1594
Qy       1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
Db       1595 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1654
Qy       1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
Db       1655 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1714
Qy       1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
Db       1715 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1774
Qy       1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
Db       1775 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1834
Qy       1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
Db       1835 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1894
Qy       1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
Db       1895 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1954
Qy       1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
Db       1955 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 2014
Qy       1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
Db       2015 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 2074
Qy       1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
Db       2075 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 2134
Qy       1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
Db       2135 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 2194
Qy       1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
Db       2195 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 2254
Qy       1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
Db       2255 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2314
Qy       1981 GAATTATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2040
Db       2315 GAATTATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2374
Qy       2041 AAAAAAAAAAAAAAA 2055
Db       2375 AAAAAAAAAAAAAAA 2389
```

RESULT 2
US-10-098-841-271
; Sequence 271, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing

```
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2
;  CURRENT APPLICATION NUMBER: US/10/098,841
;  CURRENT FILING DATE:  2002-03-13
;  PRIOR APPLICATION NUMBER: 09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 331
;  SOFTWARE: pt_PL_genes Version 1.0
; SEQ ID NO 271
;   LENGTH: 2092
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (139)..(1215)
US-10-098-841-271

  Query Match             97.0%;  Score 1997.2;  DB 14;  Length 2092;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1999;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 150

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 210

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 270

Qy        181 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 330

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        331 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 390

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        391 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAGGGTGTACCCCGTCCAGGAAGCC 450

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        451 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 510

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 570

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 630

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 690

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 750

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 810

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        811 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 870

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        871 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 930

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        931 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 990

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        991 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1050

Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1051 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1110

Qy       1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1111 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1170

Qy       1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1171 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1230

Qy       1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1231 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1290

Qy       1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1291 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1350

Qy       1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1351 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1410

Qy       1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1411 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1470

Qy       1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1471 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1530

Qy       1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1531 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1590

Qy       1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1591 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1650

Qy       1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1710

Qy       1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1711 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1770

Qy       1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1771 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1830

Qy       1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1831 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1890

Qy       1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1891 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1950

Qy       1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1951 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 2010

Qy       1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2070

Qy       1981 GAATTATGAAATAAAAAAAAAA 2002
             ||||||||||||||||||||||
Db       2071 GAATTATGAAATAAAAAAAAAA 2092


RESULT 3
US-09-925-301-260
; Sequence 260, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 260
;   LENGTH: 2048
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (66)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (67)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-260

  Query Match             85.9%;  Score 1767.8;  DB 9;  Length 2048;
  Best Local Similarity   95.9%;  Pred. No. 0;
  Matches 1882;  Conservative    4;  Mismatches    1;  Indels   76;  Gaps    4;

Qy         49 ATGCGAGCCACCCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ATGCGAGCCACCCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 215

Qy        109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 275

Qy        169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
              ||||||||||||||||||||||||||||||||
Db        276 CCCAGACTGCCCCCCTGCCTGTTGCCCCT------------------------------ 304

Qy        229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
                                                  ||||||||||||||||||
Db        305 ------------------------------------------GAGCCCGAGGAGGGCGGG 322

Qy        289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
              |||||||||| || ||||||||||||||||||||||||||||||||||| |||||||||
Db        323 CGGGCCTACCAGG-CCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 380

Qy        349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
              |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        381 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTAYGCGCGGCTGCCCCCGCACAAGCATGTG 440

Qy        409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 500

Qy        469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        501 CATGGGGACATGCACAGCCTGGTGCGAAG-CGCCACCGTATCCCTGAGCCTGAGGCTGCC 559

Qy        529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 619

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 679

Qy        649 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 739

Qy        709 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 768
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        740 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCAMGGGCCTCATACTCGGGCAAG 799

Qy        769 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 859

Qy        829 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 919

Qy        889 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 979

Qy        949 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1039

Qy       1009 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1068
              ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 CCCTTAGCYCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1099

Qy       1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGAC 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGAC 1159

Qy       1129 CACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTT 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1160 CACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTT 1219

Qy       1189 CTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGC 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1220 CTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGC 1279

Qy       1249 CTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTG 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1280 CTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTG 1339

Qy       1309 GGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGA 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1340 GGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGA 1399

Qy       1369 GACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTC 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1400 GACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTC 1459

Qy       1429 CATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGA 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1460 CATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGA 1519

Qy       1489 TGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCT 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1520 TGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCT 1579

Qy       1549 TTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1580 TTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 1639

Qy       1609 TCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCT 1668
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1640 TCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTRTTCCCAGCATCTCT 1699

Qy       1669 GTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAA 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1700 GTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAA 1759

Qy       1729 GAATCCAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCA 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1760 GAATCCAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCA 1819

Qy       1789 GGCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAG 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1820 GGCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAG 1879

Qy       1849 GGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGT 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1880 GGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGT 1939

Qy       1909 CCACAATCCCAGGTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTG 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1940 CCACAATCCCAGGTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTG 1999

Qy       1969 CCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAAA 2011
             |||||||||||||||||||||||||||||||||||||||||||
Db       2000 CCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAAA 2042
```

```
RESULT 4
US-09-799-875-9
; Sequence 9, Application US/09799875
; Patent No. US20020034780A1
; GENERAL INFORMATION:
;  APPLICANT:  Meyers, Rachel
;  APPLICANT:  Kapeller-Libermann, Rosana
;  APPLICANT:  Williamson, Mark
;  TITLE OF INVENTION:  No. US20020034780A1el Human Protein Kinases and Uses
;  TITLE OF INVENTION:  Therefor
;  FILE REFERENCE:  35800/209996
;  CURRENT APPLICATION NUMBER: US/09/799,875
;  CURRENT FILING DATE:  2001-03-06
;  PRIOR APPLICATION NUMBER:  60/182,059
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  09/659,287
;  PRIOR FILING DATE:  2000-09-12
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 1074
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-799-875-9

  Query Match             52.0%;  Score 1070.8;  DB 9;  Length 1074;
```

```
  Best Local Similarity   99.8%;  Pred. No. 1.4e-221;
  Matches 1072;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy        109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360

Qy        409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Qy        649 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Qy        709 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Qy        769 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Qy        829 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840

Qy        889 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900

Qy        949 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960

Qy       1009 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTTAGCTCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020

Qy       1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
```

RESULT 5
US-10-024-828-3
; Sequence 3, Application US/10024828
; Publication No. US20030036051A1
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase
; TITLE OF INVENTION: Punctions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/10/024,828
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/509,902A
; PRIOR FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1085
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-024-828-3

Query Match 48.8%; Score 1004.4; DB 15; Length 1085;
Best Local Similarity 99.9%; Pred. No. 3.1e-207;
Matches 1005; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 139

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 199

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 259

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 319

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        320 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 379

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 439

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 499

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 559

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 619

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 679

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 739

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 799

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 859

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 919

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTGCCTGCAGGCCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTGCCTGCAGGCCTC 979

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1039

Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1006
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1085
```

RESULT 6
US-09-919-580-580
; Sequence 580, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 580
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-580

Query Match 25.9%; Score 533; DB 10; Length 541;
Best Local Similarity 99.1%; Pred. No. 1.9e-105;
Matches 536; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGAC 1128
              || |  | ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCGCGTCGACGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGAC 60

Qy       1129 CACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTT 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTT 120

Qy       1189 CTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGC 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGC 180

Qy       1249 CTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTG 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTG 240

Qy       1309 GGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGA 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGA 300

Qy       1369 GACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTC 1428
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTC 360

Qy       1429 CATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGA 1488
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGA 420

Qy       1489 TGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCT 1548
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCT 480

Qy       1549 TTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 1608
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 540

Qy       1609 T 1609
                |
Db        541 T 541


RESULT 7
US-09-867-701-10282/c
; Sequence 10282, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10282
;   LENGTH: 426
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-10282

  Query Match             20.7%;  Score 426;  DB 10;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-82;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1567 CAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGAC 1626
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGAC 367

Qy       1627 CCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGA 1686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGA 307

Qy       1687 TTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGC 1746
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGC 247

Qy       1747 TAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGG 1806
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGG 187

Qy       1807 CTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCAC 1866
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCAC 127

Qy       1867 CTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCAT 1926
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCAT 67

Qy       1927 ACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTA 1986
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 ACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTA 7

Qy       1987 TGAAAT 1992
                ||||||
Db           6 TGAAAT 1


RESULT 8
US-09-803-719-519
; Sequence 519, Application US/09803719
; Publication No. US20030044783A1
; GENERAL INFORMATION:
;  APPLICANT: Williams, Lewis T.
;  APPLICANT:  Escobedo, Jaime
;  APPLICANT:  Innis, Michael A.
;  APPLICANT:  Garcia, Pablo Dominiguez
;  APPLICANT:  Sudduth-Klinger, Julie
;  APPLICANT:  Reinhard, Christoph
;  APPLICANT:  Giese, Klaus
;  APPLICANT:  Randazzo, Filippo
;  APPLICANT:  Kennedy, Giulia C.
;  APPLICANT:  Pot, David
;  APPLICANT:  Kassam, Altaf
;  APPLICANT:  Lamson, George
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Jones, Lee William
;  APPLICANT:  Stache-Crain, Birgit
;  TITLE OF INVENTION: Human Genes and Gene Products
;  FILE REFERENCE: 1624.002
;  CURRENT APPLICATION NUMBER: US/09/803,719
;  CURRENT FILING DATE:  2001-03-09
;  PRIOR APPLICATION NUMBER: 60/188,609
;  PRIOR FILING DATE: 2000-03-09
;  NUMBER OF SEQ ID NOS: 2396
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 519
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-803-719-519

  Query Match             16.0%;  Score 330.2;  DB 11;  Length 396;
  Best Local Similarity   97.7%;  Pred. No. 1.1e-61;
  Matches  335;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 112

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
                |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        113 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGATGGAGTTGGATGACAAC 172

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
                |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TTAGATACCGAGCGTTCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 232

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 292

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
                |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        293 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGTCCTACCAG 352

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGT 343
                |||||||||||||||||||||||    ||||||||||||||
Db        353 GCCCTGCACTGCCCTACAGGCACTTGAGATACCTGCAAGGTGT 395
```

```
RESULT 9
US-09-873-367C-341/c
; Sequence 341, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 341
;   LENGTH: 353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-341

  Query Match             15.3%;  Score 315.4;  DB 13;  Length 353;
  Best Local Similarity   95.2%;  Pred. No. 1.7e-58;
  Matches  336;  Conservative    0;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       1644 TGGGAACTGTGTT-CCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCT 1702
             ||||||||||||| |||||||||||||||||||||||||||||||||| | ||||
Db        353 TGGGAACTGTGTTCCCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCTTTCCAGCCT 294

Qy       1703 AAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAGTTCTTGTCTAACT 1762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAGTTCTTGTCTAACT 234

Qy       1763 CAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGGCTTCTGACCTGAGCAC 1822
             |||||||||||||||||||| | | |    | |    ||||||||||||||||||||
Db        233 CAAGACTGTTCTGGAATGAGGGCCAGGCCTGTCAACCATGGGGGCTTCTGACCTGAGCAC 174

Qy       1823 CAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGT 1882
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 CAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGT 114

Qy       1883 GGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACTCTAGGTTTTGGAT 1942
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 GGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACTCTAGGTTTTGGAT 54

Qy       1943 ACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTATGAAATAAA 1995
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          53 ACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTATGAAATAAA 1

RESULT 10
US-09-925-301-525
; Sequence 525, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 525
;   LENGTH: 562
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (515)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (526)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (557)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-525

  Query Match             13.1%;  Score 268.8;  DB 9;  Length 562;
  Best Local Similarity   96.7%;  Pred. No. 2.5e-48;
  Matches  267;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 333

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 393

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 453

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        454 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGSCACT 513

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCC 276
              | |||||||:|| ||:||||||||||:||:||  ||
Db        514 GNCTCCCGTYTTNGGSCCTATGTCCTYCTKGAAGCC 549

RESULT 11
US-10-291-172-116
; Sequence 116, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-045
;  CURRENT APPLICATION NUMBER: US/10/291,172
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 09/693,267
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/665,363
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: 09/616,847
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 09/596,193
;  PRIOR FILING DATE: 2000-06-17
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 116
;   LENGTH: 1909
;   TYPE: DNA
```

```
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (135)..(941)
US-10-291-172-116

  Query Match             12.8%;  Score 263.2;  DB 12;  Length 1909;
  Best Local Similarity   60.7%;  Pred. No. 6.8e-47;
  Matches  447;  Conservative    0;  Mismatches  288;  Indels    1;  Gaps    1;

Qy        258 CTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTAC 317
Db        103 CTACCTGCTGCTGCCCCTAGCCGAGCGCGAGCATGTGTCCCGGGCGCTGTGCATCCACAC 162

Qy        318 AGGCACTGAGTATAC-CTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGC 376
Db        163 TGGGACGCCGAGTGCGCTGCAAGGTGTTTCCCATTAAACACTACCAGGACAAAATCAGGC 222

Qy        377 CCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTA 436
Db        223 CTTACATCCAGCTGCCATCGCACAGCAACATTACTGGCATTGTGGAAGTGATCCTTGGGG 282

Qy        437 CCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAA 496
Db        283 AAACCAAGGCCTATGTCTTCTTTGAGAAGGACTTTGGGGACATGCACTCCTATGTGCGAA 342

Qy        497 GCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCC 556
Db        343 GCCGGAAGAGGCTGCGGGAAGAGGAAGCCGCCCGGCTCTTCAAGCAGATTGTCTCCGCCG 402

Qy        557 TGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCT 616
Db        403 TCGCCCACTGCCACCAGTCAGCCATCGTGCTGGGGGACCTGAAGCTTAGGAAGTTCGTCT 462

Qy        617 TCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGA 676
Db        463 TCTCCACGGAGGAGAGAACCCAGCTTAGACTAGAAAGTCTAGAAGACACACACATAATGA 522

Qy        677 CTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGA 736
Db        523 AGGGGGAAGATGATGCTTTGTCAGACAAACATGGCTGCCCAGCCTACGTGAGCCCTGAGA 582

Qy        737 TACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGG 796
Db        583 TCCTCAACACCACTGGGACCTACTCCGGAAAGGCTGCGGACGTTTGGAGCCTGGGGGTGA 642

Qy        797 CGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCT 856
Db        643 TGCTCTACACCCTTCTGGTTGGACGATACCCCTTCCATGACTCAGACCCCAGTGCCCTTT 702

Qy        857 TCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCT 916
Db        703 TCTCCAAAATTCGGCGTGGACAGTTCTGCATTCCTGAGCACATTTCCCCCAAAGCCAGGT 762

Qy        917 GTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCC 976
Db        763 GCCTCATTCGCAGCCTCTTGAGACGGGAGCCCTCCGAGAGACTCACTGCCCCCGAGATCC 822

Qy        977 TCCTGCACCCCTGGCT 992
Db        823 TACTGCACCCCTGGTT 838

RESULT 12
US-10-119-926-73
; Sequence 73, Application US/10119926
; Publication No. US20030104413A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030104413A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 789CIP2BCON
;  CURRENT APPLICATION NUMBER: US/10/119,926
;  CURRENT FILING DATE:  2002-04-09
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 108
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 73
;   LENGTH: 1909
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (135)..(944)
US-10-119-926-73

  Query Match             12.8%;  Score 263.2;  DB 15;  Length 1909;
  Best Local Similarity   60.7%;  Pred. No. 6.8e-47;
  Matches  447;  Conservative    0;  Mismatches  288;  Indels    1;  Gaps    1;

Qy        258 CTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTAC 317
Db        103 CTACCTGCTGCTGCCCCTAGCCGAGCGCGAGCATGTGTCCCGGGCGCTGTGCATCCACAC 162

Qy        318 AGGCACTGAGTATAC-CTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGC 376
Db        163 TGGGACGCCGAGTGCGCTGCAAGGTGTTTCCCATTAAACACTACCAGGACAAAATCAGGC 222

Qy        377 CCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTA 436
Db        223 CTTACATCCAGCTGCCATCGCACAGCAACATTACTGGCATTGTGGAAGTGATCCTTGGGG 282

Qy        437 CCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAA 496
Db        283 AAACCAAGGCCTATGTCTTCTTTGAGAAGGACTTTGGGGACATGCACTCCTATGTGCGAA 342

Qy        497 GCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCC 556
Db        343 GCCGGAAGAGGCTGCGGGAAGAGGAAGCCGCCCGGCTCTTCAAGCAGATTGTCTCCGCCG 402

Qy        557 TGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCT 616
Db        403 TCGCCCACTGCCACCAGTCAGCCATCGTGCTGGGGGACCTGAAGCTTAGGAAGTTCGTCT 462

Qy        617 TCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGA 676
Db        463 TCTCCACGGAGGAGAGAACCCAGCTTAGACTAGAAAGTCTAGAAGACACACACATAATGA 522

Qy        677 CTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGA 736
Db        523 AGGGGGAAGATGATGCTTTGTCAGACAAACATGGCTGCCCAGCCTACGTGAGCCCTGAGA 582

Qy        737 TACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGG 796
Db        583 TCCTCAACACCACTGGGACCTACTCCGGAAAGGCTGCGGACGTTTGGAGCCTGGGGGTGA 642

Qy        797 CGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCT 856
Db        643 TGCTCTACACCCTTCTGGTTGGACGATACCCCTTCCATGACTCAGACCCCAGTGCCCTTT 702

Qy        857 TCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCT 916
```

```
Db        703 TCTCCAAAATTCGGCGTGGACAGTTCTGCATTCCTGAGCACATTTCCCCCAAAGCCAGGT 762

Qy        917 GTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCC 976

Db        763 GCCTCATTCGCAGCCTCTTGAGACGGGAGCCCTCCGAGAGACTCACTGCCCCCGAGATCC 822

Qy        977 TCCTGCACCCCTGGCT 992

Db        823 TACTGCACCCCTGGTT 838


RESULT 13
US-10-084-817-300
; Sequence 300, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
;  APPLICANT: Susan Stuart
;  APPLICANT:  Jed G. Nuchtern
;  APPLICANT:  Sharon E. Plon
;  APPLICANT:  Jason M. Shohet
;  TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
;  FILE REFERENCE: PA-0046 US
;  CURRENT APPLICATION NUMBER: US/10/084,817
;  CURRENT FILING DATE:  2002-02-25
;  PRIOR APPLICATION NUMBER: 60/270,784
;  PRIOR FILING DATE: 2001-02-23
;  NUMBER OF SEQ ID NOS: 365
;  SOFTWARE: PERL Program
; SEQ ID NO 300
;   LENGTH: 4336
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030119009A1 234427.4
US-10-084-817-300

  Query Match            12.3%;  Score 253;  DB 15;  Length 4336;
  Best Local Similarity   55.7%;  Pred. No. 1.5e-44;
  Matches  484;  Conservative    0;  Mismatches  385;  Indels    0;  Gaps    0;

Qy        195 CCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGG 254

Db       1459 CCTCGGCTCCCCGAGCCCGCCCGAGACTCCGAACTTGTCGCATTGCGTTTCTTGTATCGG 1518

Qy        255 GCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCC 314

Db       1519 GAAATACTTATTGTTGGAACCTCTGGAGGGAGACCACGTTTTCGTGCCGTGCATCTGCA 1578

Qy        315 TACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGA 374

Db       1579 CAGCGGAGAGGAGCTGGTGTGCAAGGTGTTTGATATCAGCTGCTACCAGGAATCCCTGGC 1638

Qy        375 GCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGG 434

Db       1639 ACCGTGCTTTTGCCTGTCTGCTCATAGTAACATCAACCAAATCACTGAAATTATCCTGGG 1698

Qy        435 TACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCG 494

Db       1699 TGAGACCAAAGCCTATGTGTTCTTTGAGCGAAGCTATGGGGACATGCATTCCTTCGTCCG 1758

Qy        495 AAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGC 554

Db       1759 CACCTGCAAGAAGCTGAGAGAGGAGGAGGCAGCCAGACTGTTCTACCAGATTGCCTCGGC 1818

Qy        555 CCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGT 614

Db       1819 AGTGGCCCACTGCCATGACGGGGGGCTGGTGCTGCGGGACCTCAAGCTGCGGAAATTCAT 1878

Qy        615 CTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCT 674

Db       1879 CTTTAAGGACGAAGAGAGGACTCGGGTCAAGCTGGAAAGCCTGGAAGACGCCTACATTCT 1938

Qy        675 GACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGA 734

Db       1939 GCGGGGAGATGATGATTCCCTCTCCGACAAGCATGGCTGCCCGGCTTACGTAAGCCCAGA 1998

Qy        735 GATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGT 794

Db       1999 GATCTTGAACACCAGTGGCAGCTACTCGGGCAAAGCAGCCGACGTGTGGAGCCTGGGGGT 2058

Qy        795 GGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCT 854

Db       2059 GATGCTGTACACCATGTTGGTGGGGCGGTACCCTTTCCATGACATTGAACCCAGCTCCCT 2118

Qy        855 CTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCG 914

Db       2119 CTTCAGCAAGATCCGGCGTGGCCAGTTCAACATTCCAGAGACTCTGTCGCCCAAGGCCAA 2178

Qy        915 CTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCAT 974

Db       2179 GTGCCTCATCCGAAGCATTCTGCGTCGGGAGCCCTCAGAGCGGCTGACCTCGCAGGAAAT 2238

Qy        975 CCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCT 1034

Db       2239 TCTGGACCATCCTTGGTTTTCTACAGATTTTAGCGTCTCGAATTCAGCATATGGTGCTAA 2298

Qy       1035 CTGGGAGGCTGCCCAGGTGGTCCCTGATG 1063

Db       2299 GGAAGTGTCTGACCAGCTGGTGCCGGACG 2327


RESULT 14
US-10-044-090-269
; Sequence 269, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 269
;   LENGTH: 3280
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1820904CB1
US-10-044-090-269

  Query Match            12.2%;  Score 251.4;  DB 14;  Length 3280;
  Best Local Similarity   55.6%;  Pred. No. 3e-44;
  Matches  483;  Conservative    0;  Mismatches  386;  Indels    0;  Gaps    0;

Qy        195 CCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGG 254

Db        403 CCTCGGCTCCCCGAGCCCGCCCGAGACTCCGAACTTGTCGCATTGCGTTTCTTGTATCGG 462

Qy        255 GCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCC 314

Db        463 GAAATACTTATTGTTGGAACCTCTGGAGGGAGACCACGTTTTTCGTGCCGTGCATCTGCA 522

Qy        315 TACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGA 374

Db        523 CAGCGGAGAGGAGCTGGTGTGCAAGGTGTTTGATATCAGCTGCTACCAGGAATCCCTGGC 582

Qy        375 GCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGG 434

Db        583 ACCGTGCTTTTGCCTGTCTGCTCATAGTAACATCAACCAAATCACTGAAATTATCCTGGG 642

Qy        435 TACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCG 494
```

```
Db        643 TGAGACCAAAGCCTATGTGTTCTTTGAGCGAACGTATGGGGACATGCATTCCTTCGTCCG 702

Qy        495 AAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGC 554

Db        703 CACCTGCAAGAAGCTGAGAGAGGAGGAGGCAGCCAGACTGTTCTACCAGATTGCCTCGGC 762

Qy        555 CCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGT 614

Db        763 AGTGGCCCACTGCCATGACGGGGGGCTGGTGCTGCGGGACCTCAAGCTGCGGAAATTCAT 822

Qy        615 CTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCT 674

Db        823 CTTTAAGGACGAAGAGAGGACTCGGGTCAAGCTGGAAAGCCTGGAAGACGCCTACATTCT 882

Qy        675 GACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGA 734

Db        883 GCGGGGAGATGATGATTCCCTCTCCGACAAGCATGGCTGCCCGGCTTACGTAAGCCCAGA 942

Qy        735 GATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGT 794

Db        943 GATCTTGAACACCAGTGGCAGCTACTCGGGCAAAGCAGCCGACGTGTGGAGCCTGGGGGT 1002

Qy        795 GGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCT 854

Db       1003 GATGCTGTACACCATGTTGGTGGGGCGGTACCCTTTCCATGACATTGAACCCAGCTCCCT 1062

Qy        855 CTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCG 914

Db       1063 CTTCAGCAAGACCCGGCGTGGCCAGTTCAACATTCCAGAGACTCTGTCGCCCAAGGCCAA 1122

Qy        915 CTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCAT 974

Db       1123 GTGCCTCATCCGAAGCATTCTGCGTCGGGAGCCCTCAGAGCGGCTGACCTCGCAGGAAAT 1182

Qy        975 CCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCT 1034

Db       1183 TCTGGACCATCCTTGGTTTTCTACAGATTTTAGCGTCTCGAATTCAGCATATGGTGCTAA 1242

Qy       1035 CTGGGAGGCTGCCCAGGTGGTCCCTGATG 1063

Db       1243 GGAAGTGTCTGACCAGCTGGTGCCGGACG 1271


RESULT 15
US-10-240-965-168
; Sequence 168, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  SHIFFMAN, Dov
;  APPLICANT:  SOMOGYI, Roland
;  APPLICANT:  LAWN, Richard M.
;  APPLICANT:  SEILHAMER, Jeffrey J.
;  APPLICANT:  PORTER, Gordon J.
;  APPLICANT:  MIKITA, Thomas
;  APPLICANT:  TAI, Julie
;  TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
;  FILE REFERENCE: PA-0025 PCT
;  CURRENT APPLICATION NUMBER: US/10/240,965
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/195,106
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 276
;  SOFTWARE: PERL Program
; SEQ ID NO 168
;   LENGTH: 3324
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030165924A1 197301.4
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 254, 378
;   OTHER INFORMATION: a, t, c, g, or other
US-10-240-965-168

  Query Match             12.0%;  Score 247;  DB 13;  Length 3324;
  Best Local Similarity   61.1%;  Pred. No. 2.7e-43;
  Matches  400;  Conservative    0;  Mismatches  255;  Indels    0;  Gaps    0;

Qy        338 AGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGC 397

Db        635 AGGTGTTTCCCATTAAACACTACCAGGACAAAATCAGGCCTTACATCCAGCTGCCATCGC 694

Qy        398 ACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTT 457

Db        695 ACAGCAACATTACTGGCATTGTGGAAGTGATCCTTGGGGAAACCAAGGCCTATGTCTTCT 754

Qy        458 TCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGC 517

Db        755 TTGAGAAGGACTTTGGGGACATGCACTCCTATGTGCGAAGCCGGAAGAGGCTGCGGGAAG 814

Qy        518 CTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACG 577

Db        815 AGGAAGCCGCCCGGCTCTTCAAGCAGATTGTCTCCGCCGTCGCCCACTGCCACCAGTCAG 874

Qy        578 GTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGA 637

Db        875 CCATCGTGCTGGGGGACCTGAAGCTTAGGAAGTTCGTCTTCTCCACGGAGGAGAGAACCC 934

Qy        638 AGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT 697

Db        935 AGCTTAGACTAGAAAGTCTAGAAGACACACACATAATGAAGGGGGAAGATGATGCTTTGT 994

Qy        698 GGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCAT 757

Db        995 CAGACAAACATGGCTGCCCAGCCTACGTGAGCCCTGAGATCCTCAACACCACTGGGACCT 1054

Qy        758 ACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCG 817

Db       1055 ACTCCGGAAAGGCTGCGGACGTTTGGAGCCTGGGGGTGATGCTCTACACCCTTCTGGTTG 1114

Qy        818 GCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGG 877

Db       1115 GACGATACCCCTTCCATGACTCAGACCCCAGTGCCCTTTTCTCCAAAATTCGGCGTGGAC 1174

Qy        878 CCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTC 937

Db       1175 AGTTCTGCATTCCTGAGCACATTTCCCCCAAAGCCAGGTGCCTCATTCGCAGCCTCTTGA 1234

Qy        938 GTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCT 992

Db       1235 GACGGGAGCCCTCCGAGAGACTCACTGCCCCCGAGATCCTACTGCACCCCTGGTT 1289


Search completed: January 16, 2004, 20:19:03
Job time : 717.201 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:03:57 ; Search time 21 Seconds
(without alignments)
721.299 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 358
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	234	65.4	360	4	US-09-509-902A-9	Sequence 9, Appli
2	8	2.2	731	4	US-09-252-991A-25890	Sequence 25890, A
3	8	2.2	778	4	US-09-252-991A-18605	Sequence 18605, A
4	7	2.0	67	4	US-09-252-991A-31777	Sequence 31777, A
5	7	2.0	73	4	US-09-328-352-5099	Sequence 5099, Ap
6	7	2.0	90	4	US-09-252-991A-20965	Sequence 20965, A
7	7	2.0	133	4	US-09-384-162-16	Sequence 16, Appl
8	7	2.0	146	4	US-09-252-991A-24993	Sequence 24993, A
9	7	2.0	153	4	US-09-252-991A-17117	Sequence 17117, A
10	7	2.0	160	4	US-09-482-273-211	Sequence 211, App
11	7	2.0	163	2	US-08-469-537A-83	Sequence 83, Appl
12	7	2.0	167	4	US-09-252-991A-23665	Sequence 23665, A
13	7	2.0	190	2	US-08-469-537A-35	Sequence 35, Appl
14	7	2.0	190	2	US-08-469-537A-36	Sequence 36, Appl
15	7	2.0	201	1	US-08-444-083-8	Sequence 8, Appli
16	7	2.0	201	1	US-08-286-304-8	Sequence 8, Appli
17	7	2.0	201	1	US-08-442-745-8	Sequence 8, Appli
18	7	2.0	201	1	US-08-443-129-8	Sequence 8, Appli
19	7	2.0	201	1	US-08-443-952-8	Sequence 8, Appli
20	7	2.0	201	1	US-08-443-130-8	Sequence 8, Appli
21	7	2.0	201	1	US-08-792-019B-11	Sequence 11, Appl
22	7	2.0	201	3	US-09-106-182-4	Sequence 4, Appli
23	7	2.0	201	3	US-08-988-819-11	Sequence 11, Appl
24	7	2.0	201	3	US-08-898-911-8	Sequence 8, Appli
25	7	2.0	201	3	US-09-016-534-11	Sequence 11, Appl
26	7	2.0	201	4	US-09-648-183-3	Sequence 3, Appli
27	7	2.0	201	5	PCT-US95-04467-8	Sequence 8, Appli
28	7	2.0	222	4	US-09-384-162-8	Sequence 8, Appli
29	7	2.0	232	4	US-09-149-476-623	Sequence 623, App
30	7	2.0	279	2	US-08-469-537A-51	Sequence 51, Appl
31	7	2.0	285	2	US-08-469-537A-73	Sequence 73, Appl
32	7	2.0	286	4	US-09-252-991A-27746	Sequence 27746, A
33	7	2.0	324	4	US-09-252-991A-19999	Sequence 19999, A
34	7	2.0	346	3	US-08-888-429A-18	Sequence 18, Appl
35	7	2.0	374	2	US-08-958-240-2	Sequence 2, Appli
36	7	2.0	380	3	US-08-888-429A-30	Sequence 30, Appl
37	7	2.0	380	4	US-09-252-991A-22319	Sequence 22319, A
38	7	2.0	382	4	US-09-252-991A-18552	Sequence 18552, A
39	7	2.0	389	3	US-08-888-429A-20	Sequence 20, Appl
40	7	2.0	393	3	US-08-888-429A-21	Sequence 21, Appl
41	7	2.0	419	3	US-08-888-429A-28	Sequence 28, Appl
42	7	2.0	419	4	US-09-446-754-6	Sequence 6, Appli
43	7	2.0	419	4	US-09-206-166-2	Sequence 2, Appli
44	7	2.0	419	4	US-09-206-166-5	Sequence 5, Appli
45	7	2.0	435	4	US-09-446-754-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-509-902A-9
; Sequence 9, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-902A-9

Query Match 65.4%; Score 234; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 2.4e-221;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDM 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         127 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDM 186

Qy         145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         187 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 246

Qy         205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 264
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         247 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 306

Qy         265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD 318
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         307 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD 360
```

RESULT 2
US-09-252-991A-25890
; Sequence 25890, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 25890
;   LENGTH: 731
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25890

  Query Match             2.2%;  Score 8;  DB 4;  Length 731;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TPLAAPAG 11
              ||||||||
Db        306 TPLAAPAG 313


RESULT 3
US-09-252-991A-18605
; Sequence 18605, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 18605
;   LENGTH: 778
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18605

  Query Match             2.2%;  Score 8;  DB 4;  Length 778;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        339 GLGLDEAR 346
              ||||||||
Db         78 GLGLDEAR 85


RESULT 4
US-09-252-991A-31777
; Sequence 31777, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 31777
;   LENGTH: 67
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31777

  Query Match             2.0%;  Score 7;  DB 4;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        177 GLVLRDL 183
              |||||||
Db         16 GLVLRDL 22


RESULT 5
US-09-328-352-5099
; Sequence 5099, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5099
;   LENGTH: 73
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-5099

  Query Match             2.0%;  Score 7;  DB 4;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        209 LTGPDDS 215
              |||||||
Db         10 LTGPDDS 16


RESULT 6
US-09-252-991A-20965
; Sequence 20965, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 20965
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20965

  Query Match             2.0%;  Score 7;  DB 4;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 PAGSLSR 15
              |||||||
Db         30 PAGSLSR 36


RESULT 7
US-09-384-162-16
; Sequence 16, Application US/09384162
```

```
; Patent No. 6376747
; GENERAL INFORMATION:
;  APPLICANT: Xing, Ti
;  APPLICANT: Malik, Kamal
;  APPLICANT: Martin-Heller, Teresa
;  APPLICANT: Miki L., Brian
;  TITLE OF INVENTION: No. 6376747el Plant-Derived Map Kinase Kinase
;  FILE REFERENCE: 08-884280US
;  CURRENT APPLICATION NUMBER: US/09/384,162
;  CURRENT FILING DATE: 1999-08-27
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
;   LENGTH: 133
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-384-162-16

  Query Match             2.0%;  Score 7;  DB 4;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        243 DVWSLGV 249
              |||||||
Db         69 DVWSLGV 75


RESULT 8
US-09-252-991A-24993
; Sequence 24993, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 24993
;   LENGTH: 146
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24993

  Query Match             2.0%;  Score 7;  DB 4;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         47 LLPLSPP 53
              |||||||
Db         63 LLPLSPP 69


RESULT 9
US-09-252-991A-17117
; Sequence 17117, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 17117
;   LENGTH: 153
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17117

  Query Match             2.0%;  Score 7;  DB 4;  Length 153;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 PLAAPAG 11
              |||||||
Db        109 PLAAPAG 115


RESULT 10
US-09-482-273-211
; Sequence 211, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 71 Human Secreted Proteins
;  FILE REFERENCE: PZ030P1
;  CURRENT APPLICATION NUMBER: US/09/482,273
;  CURRENT FILING DATE: 2000-01-13
;  EARLIER APPLICATION NUMBER: PCT/US99/15849
;  EARLIER FILING DATE: 1999-07-14
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/092,922
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/092,956
;  EARLIER FILING DATE: 1998-07-15
;  NUMBER OF SEQ ID NOS: 267
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-482-273-211

  Query Match             2.0%;  Score 7;  DB 4;  Length 160;
  Best Local Similarity   100.0%;  Pred. No. 57;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         44 PPCLLPL 50
              |||||||
Db          3 PPCLLPL 9


RESULT 11
US-08-469-537A-83
; Sequence 83, Application US/08469537A
; Patent No. 5843749
;  GENERAL INFORMATION:
;    APPLICANT:  Maisonpierre, et al.
;    TITLE OF INVENTION:  EHK AND ROR TYROSINE
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Regeneron Pharmaceuticals, Inc.
;      STREET:  777 Old Saw Mill River Road
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  U.S.A.
;      ZIP:  10591
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/469,537A
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  USSN 08/406,247
;       FILING DATE:  17-MAR-1995
;       APPLICATION NUMBER:  USSN 08/144,992
;       FILING DATE:  28-OCT-1993
;       APPLICATION NUMBER:  USSN 07/736,559
;       FILING DATE:  26-JUL-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kempler, Ph.D., Gail M
;       REGISTRATION NUMBER:  32,143
;       REFERENCE/DOCKET NUMBER:  REG 070C
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  914-345-7400
;       TELEFAX:  914-345-7721
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  83:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  163 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  peptide
US-08-469-537A-83

  Query Match             2.0%;  Score 7;  DB 2;  Length 163;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        243 DVWSLGV 249
              |||||||
Db         90 DVWSLGV 96


RESULT 12
US-09-252-991A-23665
; Sequence 23665, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 23665
;   LENGTH: 167
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23665

  Query Match             2.0%;  Score 7;  DB 4;  Length 167;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         39 PQPRLPP 45
              |||||||
Db        112 PQPRLPP 118


RESULT 13
US-08-469-537A-35
; Sequence 35, Application US/08469537A
; Patent No. 5843749
;  GENERAL INFORMATION:
;    APPLICANT:  Maisonpierre, et al.
;    TITLE OF INVENTION:  EHK AND ROR TYROSINE
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Regeneron Pharmaceuticals, Inc.
;      STREET:  777 Old Saw Mill River Road
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  U.S.A.
;      ZIP:  10591
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,537A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/406,247
;      FILING DATE:  17-MAR-1995
;      APPLICATION NUMBER:  USSN 08/144,992
;      FILING DATE:  28-OCT-1993
;      APPLICATION NUMBER:  USSN 07/736,559
;      FILING DATE:  26-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kempler, Ph.D., Gail M
;      REGISTRATION NUMBER:  32,143
;      REFERENCE/DOCKET NUMBER:  REG 070C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  914-345-7400
;      TELEFAX:  914-345-7721
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  190 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-469-537A-35

  Query Match             2.0%;  Score 7;  DB 2;  Length 190;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        243 DVWSLGV 249
              |||||||
Db        103 DVWSLGV 109


RESULT 14
US-08-469-537A-36
; Sequence 36, Application US/08469537A
; Patent No. 5843749
;  GENERAL INFORMATION:
;    APPLICANT:  Maisonpierre, et al.
;    TITLE OF INVENTION:  EHK AND ROR TYROSINE
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Regeneron Pharmaceuticals, Inc.
;      STREET:  777 Old Saw Mill River Road
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  U.S.A.
;      ZIP:  10591
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,537A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/406,247
;      FILING DATE:  17-MAR-1995
;      APPLICATION NUMBER:  USSN 08/144,992
;      FILING DATE:  28-OCT-1993
;      APPLICATION NUMBER:  USSN 07/736,559
;      FILING DATE:  26-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kempler, Ph.D., Gail M
;      REGISTRATION NUMBER:  32,143
;      REFERENCE/DOCKET NUMBER:  REG 070C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  914-345-7400
;      TELEFAX:  914-345-7721
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  190 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-469-537A-36

  Query Match             2.0%;  Score 7;  DB 2;  Length 190;
  Best Local Similarity   100.0%;  Pred. No. 67;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        243 DVWSLGV 249
              |||||||
Db        103 DVWSLGV 109


RESULT 15
US-08-444-083-8
; Sequence 8, Application US/08444083
; Patent No. 5571675
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Joffre
;    APPLICANT:  Chien, Kenneth
;    APPLICANT:  King, Kathleen
;    APPLICANT:  Pennica, Diane
;    APPLICANT:  Wood, William
;    TITLE OF INVENTION:  Cardiac Hypertrophy Factor and Uses Therefor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,083
;      FILING DATE:  17-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/233609
;      FILING DATE:  25-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/286304
;      FILING DATE:  05-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  894P1D5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  201 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-444-083-8

  Query Match             2.0%;  Score 7;  DB 1;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        282 AGLSAPA 288
              |||||||
Db         69 AGLSAPA 75


Search completed: January 15, 2004, 15:07:22
Job time : 22 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:56:21 ; Search time 21 Seconds
(without alignments)
721.299 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 1891
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1675	88.6	360	4	US-09-509-902A-9	Sequence 9, Appli
2	290.5	15.4	607	2	US-08-878-989-15	Sequence 15, Appl
3	290.5	15.4	607	3	US-09-272-796-15	Sequence 15, Appl
4	274.5	14.5	633	3	US-08-557-006C-43	Sequence 43, Appl
5	273.5	14.5	668	4	US-09-930-181-2	Sequence 2, Appli
6	268.5	14.2	252	2	US-07-857-224B-26	Sequence 26, Appl
7	265.5	14.0	339	3	US-08-688-988-33	Sequence 33, Appl
8	263.5	13.9	264	2	US-07-857-224B-19	Sequence 19, Appl
9	263	13.9	257	2	US-07-857-224B-25	Sequence 25, Appl
10	260.5	13.8	552	3	US-08-557-006C-40	Sequence 40, Appl
11	256	13.5	354	3	US-08-688-988-29	Sequence 29, Appl
12	253	13.4	264	2	US-07-857-224B-18	Sequence 18, Appl
13	252.5	13.4	433	2	US-08-913-050A-7	Sequence 7, Appli
14	252.5	13.4	433	2	US-08-749-902-5	Sequence 5, Appli
15	248	13.1	359	3	US-08-688-988-32	Sequence 32, Appl
16	247.5	13.1	685	2	US-08-878-989-1	Sequence 1, Appli
17	247.5	13.1	685	3	US-09-136-282-2	Sequence 2, Appli
18	247.5	13.1	685	3	US-09-272-796-1	Sequence 1, Appli
19	247.5	13.1	685	3	US-09-505-744-2	Sequence 2, Appli
20	246.5	13.0	382	3	US-09-142-551A-2	Sequence 2, Appli
21	244.5	12.9	604	4	US-09-523-849-35	Sequence 35, Appl
22	243.5	12.9	556	4	US-09-800-960-4	Sequence 4, Appli
23	243.5	12.9	565	4	US-09-800-960-2	Sequence 2, Appli
24	243.5	12.9	603	4	US-09-930-181-17	Sequence 17, Appl
25	243	12.9	273	1	US-08-252-995D-10	Sequence 10, Appl
26	243	12.9	273	2	US-08-834-108-10	Sequence 10, Appl
27	243	12.9	416	1	US-08-252-995D-2	Sequence 2, Appli
28	243	12.9	416	2	US-08-834-108-2	Sequence 2, Appli
29	243	12.9	464	1	US-08-252-995D-6	Sequence 6, Appli
30	243	12.9	464	2	US-08-834-108-6	Sequence 6, Appli
31	243	12.9	925	1	US-08-252-995D-4	Sequence 4, Appli
32	243	12.9	925	2	US-08-834-108-4	Sequence 4, Appli
33	242.5	12.8	351	3	US-08-688-988-28	Sequence 28, Appl
34	242.5	12.8	353	3	US-08-688-988-31	Sequence 31, Appl
35	241	12.7	295	1	US-07-951-715A-23	Sequence 23, Appl
36	241	12.7	295	2	US-08-459-448A-23	Sequence 23, Appl
37	241	12.7	295	3	US-08-459-595A-23	Sequence 23, Appl
38	241	12.7	295	3	US-08-459-504B-23	Sequence 23, Appl
39	241	12.7	295	3	US-08-459-444-23	Sequence 23, Appl
40	241	12.7	295	4	US-09-547-422-23	Sequence 23, Appl
41	240	12.7	776	4	US-09-523-849-34	Sequence 34, Appl
42	239	12.6	345	3	US-09-101-146-1	Sequence 1, Appli
43	237	12.5	363	3	US-08-688-988-30	Sequence 30, Appl
44	235.5	12.5	463	1	US-07-951-715A-25	Sequence 25, Appl
45	235.5	12.5	463	2	US-08-459-448A-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-509-902A-9
; Sequence 9, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-902A-9

Query Match 88.6%; Score 1675; DB 4; Length 360;
Best Local Similarity 99.7%; Pred. No. 9.9e-164;
Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 102

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        103 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 162

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 222

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 282

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 342

Qy        301 AERLTATGILLHPWLRQD 318
              ||||||||||||||||||
Db        343 AERLTATGILLHPWLRQD 360
```

```
RESULT 2
US-08-878-989-15
; Sequence 15, Application US/08878989
; Patent No. 5885803
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  DISEASE ASSOCIATED PROTEIN
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/878,989
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0321 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  607 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1827450
US-08-878-989-15

  Query Match             15.4%;  Score 290.5;  DB 2;  Length 607;
  Best Local Similarity   29.2%;  Pred. No. 3.4e-21;
  Matches   92;  Conservative   45;  Mismatches  129;  Indels   49;  Gaps   12;

Qy         48 LPLSPP----TAPDRATAVATASRLGPYVLLEPEEGG--RAYRALHCPTGTEYTCKVYPV 101
              || | |    | |            ||        :||  | | |    ||: |  || |
Db          5 LPTSDPGRLITDPRSGRTYLKGRLLG--------KGGFARCYEATDTETGSAYAVKVIPQ 56

Qy        102 QEALAVLEPYARLP-----------PHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVR 149
                | :|: |               |:|: | :         :| |    :   :  : :
Db         57 SR---VAKPHQREKILNEIELHRDLQHRHIVRFSHHFEDADNIYIFLELCSRKSLAHIWK 113

Qy        150 SRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVL 209
              :|| : |||     ||: : | : || |:: |||||  |     | :| : :   :  |
Db        114 ARHTLLEPEVRYYLRQILSGLKYLHQRGILHRDLKLGNFFIT--ENMELKVGDFGLAARL 171

Qy        210 TGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPV 267
                |:     |  |  | || ||:|  :    |  |||||||  ::|:| |  ||: ::
Db        172 EPPEQR--KKTICGTPNYVAPEVLLRQG--HGPEADVWSLGCVMYTLLCGSPPFETADLK 227

Qy        268 LLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQ----DPMPLA 323
                 :  |:: | ||| || ||| |:  :||  | :| :   || | :  :    | :|:: 
Db        228 ETYRCIKQVHYTLPASLSLPARQLLAAILRASPRDRPSIDQILRHDFFTKGYTPDRLPIS 287

Qy        324 PTRSHLWEAAQVVPD 338
                      :   |||
Db        288 --------SCVTVPD 294


RESULT 3
US-09-272-796-15
; Sequence 15, Application US/09272796
; Patent No. 6207148
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  DISEASE ASSOCIATED PROTEIN
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/272,796
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/878,989
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0321 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  607 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1827450
US-09-272-796-15

  Query Match             15.4%;  Score 290.5;  DB 3;  Length 607;
  Best Local Similarity   29.2%;  Pred. No. 3.4e-21;
  Matches   92;  Conservative   45;  Mismatches  129;  Indels   49;  Gaps   12;

Qy         48 LPLSPP----TAPDRATAVATASRLGPYVLLEPEEGG--RAYRALHCPTGTEYTCKVYPV 101
              || | |    | |            ||        :||  | | |    ||: |  || |
Db          5 LPTSDPGRLITDPRSGRTYLKGRLLG--------KGGFARCYEATDTETGSAYAVKVIPQ 56

Qy        102 QEALAVLEPYARLP-----------PHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVR 149
                | :|: |               |:|: | :         :| |    :   :  : :
Db         57 SR---VAKPHQREKILNEIELHRDLQHRHIVRFSHHFEDADNIYIFLELCSRKSLAHIWK 113
```

```
Qy        150 SRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVL 209
              :|| : |||     ||: : | : || |:: ||||| |     | :| : :   :  |
Db        114 ARHTLLEPEVRYYLRQILSGLKYLHQRGILHRDLKLGNFFIT--ENMELKVGDFGLAARL 171

Qy        210 TGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPV 267
                |:     |  |  | || ||:|  :    |  ||||||  ::|:| |  ||: ::
Db        172 EPPEQR--KKTICGTPNYVAPEVLLRQG--HGPEADVWSLGCVMYTLLCGSPPFETADLK 227

Qy        268 LLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQ----DPMPLA 323
              :  |::  | ||| || ||| |:  :||  | :| :   || | :  :    | :|::
Db        228 ETYRCIKQVHYTLPASLSLPARQLLAAILRASPRDRPSIDQILRHDFFTKGYTPDRLPIS 287

Qy        324 PTRSHLWEAAQVVPD 338
                    :  |||
Db        288 --------SCVTVPD 294


RESULT 4
US-08-557-006C-43
; Sequence 43, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
;  APPLICANT:  Beri, Rajindar K.
;  APPLICANT:  Carling, David
;  APPLICANT:  Forder, Robert A.
;  TITLE OF INVENTION:  NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
;  FILE REFERENCE:  NGAP/PHM37588/UST
;  CURRENT APPLICATION NUMBER: US/08/557,006C
;  CURRENT FILING DATE:  1996-03-06
;  PRIOR APPLICATION NUMBER:  PCT/GB94/01093
;  PRIOR FILING DATE:  1994-05-20
;  PRIOR APPLICATION NUMBER:  GB 9310489.1
;  PRIOR FILING DATE:  1993-05-21
;  PRIOR APPLICATION NUMBER:  GB 9318010.7
;  PRIOR FILING DATE:  1993-08-31
;  NUMBER OF SEQ ID NOS:  44
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 43
;   LENGTH: 633
;   TYPE: PRT
;   ORGANISM: Yeast
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(633)
;   OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

  Query Match             14.5%;  Score 274.5;  DB 3;  Length 633;
  Best Local Similarity   27.8%;  Pred. No. 1.6e-19;
  Matches   88;  Conservative   53;  Mismatches  138;  Indels   37;  Gaps   11;

Qy         51 SPPTAPDRATAVATASRLGPYVLLEP-EEG--GRAYRALHCPTGTEYTCKVYPVQEALAV 107
              |  |   :  :::|  : :| | :::   ||  |:   | |  || :   |:   :: ||
Db         35 SNSTLNNPKSSLADGAHIGNYQIVKTLGEGSFGKVKLAYHTTTGQKVALKIIN-KKVLAK 93

Qy        108 LE---------PYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPE 158
               :          | ||  | |: :  :|:      :              ::   :  | :: | |
Db         94 SDMQGRIEREISYLRLLRHPHIIKLYDVIKSKDEIIMVIEYAGNELFDYIVQRDKMSEQE 153

Qy        159 AAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD 218
              |    |:|: :|: :||:| :| ||||     :  :     |:     |  | ::|    |  :
Db        154 ARRFFQQIISAVEYCHRHKIVHRDLKPENLLLDEHLNVKIADFGL--SNIMT---DGNFL 208

Qy        219 KHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG 276
              | :|  | |  ||::| :  |:|   |||| || |: ||     || |     :||  |  |
Db        209 KTSCGSPNYAAPEVISGKL-YAGPEVDVWSCGVILYVMLCRRLPFDDESIPVLFKNISNG 267

Qy        277 AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD-PMPLAPTRSHLWEAAQV 335
              | ||  ||  |  |:: :|   |  |::    |:   | : | |  | |
Db        268 VYTLPKFLSPGAAGLIKRMLIVNPLNRISIHEIMQDDWFKVDLPEYLLP---------- 316

Qy        336 VPDGLGLDEAREEEGD 351


                  ||   |    ||| :
Db        317 -PD---LKPHPEEENE 328


RESULT 5
US-09-930-181-2
; Sequence 2, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
;  APPLICANT: OriGene Technologies
;  TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
;  FILE REFERENCE: 16U 101 V1
;  CURRENT APPLICATION NUMBER: US/09/930,181
;  CURRENT FILING DATE:  2001-08-16
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 668
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-930-181-2

  Query Match             14.5%;  Score 273.5;  DB 4;  Length 668;
  Best Local Similarity   27.6%;  Pred. No. 2.2e-19;
  Matches   92;  Conservative   52;  Mismatches  138;  Indels   51;  Gaps   14;

Qy         65 ASRLGPYVL---LEPEEGGRAYRALHCPTGTEYTCKV----------YPVQEALAVLEP 110
              |  :||| |    |   : |      :|| |  :   |:              |:  :|:|
Db         13 AQYVGPYRLEKTLGKGQTGLVKLGVHCVTCQKVAIKIVNREKLSESVLMKVEREIAIL-- 70

Qy        111 YARLPPHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATA 169
                :|  | || :  :|    : ||       : |::   :  : |:   ||   |||: :|
Db         71 --KLIEHPHVLKLHDVYENKKYLYLVLEHVSGGELFDYLVKKGRLTPKEARKFFRQIISA 128

Qy        170 LAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHAC--PAYVG 227
              |  || | :  ||||    :    |:  : : :   : :  |  ||| :  :|  | |
Db        129 LDFCHSHSICHRDLKPENLLL--DEKNNIRIADFGMASLQVG--DSLLET-SCGSPHYAC 183

Qy        228 PEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAP 287
              ||::     | |: ||||| || || :| |  || |      |  |::|| : :|  :
Db        184 PEVIRGE-KYDGRKADVWSCGVILFALLVGALPFDDDNLRQLLEKVKRGVFHMPHFIPPD 242

Qy        288 ARCLVRCLLRREPAERLTATGILLHPWL---RQDPMPLAP------TRSHLWEAAQVVPD 338
              : |:| ::  :  | |||   |  | |     : :| |  |        || |     : ||
Db        243 CQSLLRGMIEVDAARRLTLEHIQKHIWYIGGKNEPEPEQPIPRKVQIRS-LPSLEDIDPD 301

Qy        339 ------GLG--------LDEAREEEGDREVVLY 357
                   ||          | :   || ::| ::|
Db        302 VLDSMHSLGCFRDRNKLLQDLLSEEENQEKMIY 334


RESULT 6
US-07-857-224B-26
; Sequence 26, Application US/07857224B
; Patent No. 5958784
;  GENERAL INFORMATION:
;    APPLICANT:  Benner, Steven A.
;    TITLE OF INVENTION:  Predicting Folded Structures of Proteins
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Steven A. Benner
;      STREET:  Hadlaubstrasse 151
;      CITY:  Zurich
;      STATE:  none
;      COUNTRY:  Switzerland
;      ZIP:  (note: this is an international post code) CH-8092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette, 1.4 Mb storage
;      COMPUTER:  Apple MacIntosh
;      OPERATING SYSTEM:  MacIntosh 7.0
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/857,224B
```

```
;      FILING DATE:  03/25/92
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:  none
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (International) 41 1 632 2830
;      TELEFAX:  (International) 41 1 262 2437
;      TELEX:  none
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  252
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Schizosaccharomyces pombe
;    FEATURE:  Protein kinase; Table 8 Column 29
;    PUBLICATION INFORMATION:
;      AUTHORS:
;      AUTHORS:  Hanks, S. K.
;      AUTHORS:  Quinn, A. M.
;      AUTHORS:  Hunter, T.
;      TITLE:  The protein kinase family
;      JOURNAL:  Science
;      VOLUME:  241
;      PAGES:  42-52
;      DATE:  1988
US-07-857-224B-26

  Query Match             14.2%;  Score 268.5;  DB 2;  Length 252;
  Best Local Similarity   31.5%;  Pred. No. 1.8e-19;
  Matches   74;  Conservative   44;  Mismatches  108;  Indels    9;  Gaps    6;

Qy          85 ALHCPTGTEYTCKVYPVQEALAVLE-PYARLPPHKHVARPTEVLAGTQLLY-AFFTRTHG 142
               | |  ||      |: |:: |   :|    ||  | :: |  :|    | :| |      |
Db          20 AKHAKTGDLAAIKIIPIRYASIGMEILMMRLLRHPNILRLYDVWTDHQHMYLALEYVPDG 79

Qy         143 DMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLEN 202
               ::   :|    : | |||    |:  |:||||:     |||||   :    |::   ::
Db          80 ELFHYIRKHGPLSEREAAHYLSQILDAVAHCHRFRVRHRDLKLENILIKVNEQQ---IKI 136

Qy         203 LEDSCVLTGPDDSLWDKHACPA--YVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYP 260
                :       |:||  : : | :  |: |||:| :  | |  ||||| || |:::|:    |
Db         137 ADFGMATVEPNDSCLENY-CGSLHYLAPEIVSHK-PYRGAPADVWSCGVILYSLLSNKLP 194

Qy         261 FQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWL 315
               |      ::: ||| ||| ||: :|: |: |:  :|   |: |:|      ||:|
Db         195 FGGQNTDVIYNKIRHGAYDLPSSISSAAQDLLHRMLDVNPSTRITIPEPFSHPFL 249


RESULT 7
US-08-688-988-33
; Sequence 33, Application US/08688988B
; Patent No. 6096545
; GENERAL INFORMATION:
;  APPLICANT: Lefebvre, Daniel D.
;  APPLICANT: Malboobi, Mohammad A.
;  TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
;  FILE REFERENCE: PPL96-03
;  CURRENT APPLICATION NUMBER: US/08/688,988B
;  CURRENT FILING DATE: 1996-07-31
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 33
;   LENGTH: 339
;   TYPE: PRT
;   ORGANISM: Glycine max
US-08-688-988-33

  Query Match             14.0%;  Score 265.5;  DB 3;  Length 339;
  Best Local Similarity   28.7%;  Pred. No. 5.5e-19;
  Matches   86;  Conservative   35;  Mismatches   86;  Indels   93;  Gaps   12;

Qy         117 HKHVARPTEVLAGTQLLYAFFTRTH----------GDMHSLVRSRHRIPEPEAAVLFRQM 166
               | :: |  ||          | | ||          |::   : :  |: | ||   |:|:
Db          58 HPNIIRFKEV---------FLTPTHLAIVLEYAAGGELFERICNAGRLSEDEARFFFQQL 108

Qy         167 ATALAHCHQHGLVLRDLKL--------------CRFVFADRERKKLVLENLEDSCVLTG 211
                : :::||   :  |||||              | | |:    |  :| :    | | |
Db         109 ISGVSYCHSMQICHRDLKLENTLLDGNPAPRLKICDFGFS----KSALLHSQPKSTVGT- 163

Qy         212 PDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF- 270
                        |||: ||:| ||  | || |||||| || |: || | |||:| |    |
Db         164 ----------PAYIAPEVL-SRKEYDGKVADVWSCGVTLYVMLVGAYPFEDPEDPKNFR 211

Qy         271 ---GKIRRGAYALP--AGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD-PMPL-- 322
                  |:|    ||:|    :|   | |:  :    ||:|:  : |  | | |:: |  :
Db         212 KSIGRIMSVQYAIPDYVRVSKECRHLISRIFVANPAKRINISEIKQHLWFRKNLPREIIE 271

Qy         323 --------------------------APTRSHLWEAAQVVPDGLGL-DEAREEEGDREV 354
                                       | |: |  | |     | |  |  | :| : ||
Db         272 AERRGYEETQKDQPSQSVEEIMQIIQEARTKIHTGEQA-----GTGTSDVVRGDEANEEV 326


RESULT 8
US-07-857-224B-19
; Sequence 19, Application US/07857224B
; Patent No. 5958784
;  GENERAL INFORMATION:
;    APPLICANT:  Benner, Steven A.
;    TITLE OF INVENTION:  Predicting Folded Structures of Proteins
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Steven A. Benner
;      STREET:  Hadlaubstrasse 151
;      CITY:  Zurich
;      STATE:  none
;      COUNTRY:  Switzerland
;      ZIP:  (note: this is an international post code) CH-8092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette, 1.4 Mb storage
;      COMPUTER:  Apple MacIntosh
;      OPERATING SYSTEM:  MacIntosh 7.0
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/857,224B
;      FILING DATE:  03/25/92
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:  none
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (International) 41 1 632 2830
;      TELEFAX:  (International) 41 1 262 2437
;      TELEX:  none
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  rat
;    FEATURE:  Protein kinase; Table 8 Column 21
;    PUBLICATION INFORMATION:
;      AUTHORS:
;      AUTHORS:  Hanks, S. K.
;      AUTHORS:  Quinn, A. M.
;      AUTHORS:  Hunter, T.
;      TITLE:  The protein kinase family
;      JOURNAL:  Science
;      VOLUME:  241
;      PAGES:  42-52
;      DATE:  1988
US-07-857-224B-19

  Query Match             13.9%;  Score 263.5;  DB 2;  Length 264;
```

```
  Best Local Similarity   30.0%;  Pred. No. 6.1e-19;
  Matches   74;  Conservative   47;  Mismatches  111;  Indels   15;  Gaps    6;

Qy          84 RALHCPTGTEYTCKVYPVQEALA----VLEPYA---RLPPHKHVARPTEVLAGTQLLYAF 136
               | :   || ||  |:   ::  |      ||  |    ||  | :: |  : ::     |
Db          19 RCVKLCTGHEYAAKIINTKKLSARDHQKLEREARICRLLKHSNIVRLHDSISEEGFHYLV 78

Qy         137 FTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRER 195
               |   | |::   : :|    | :|:   :|:  |: |||| |:| ||||    : | : :
Db          79 FDLVTGGELFEDIVAREYYSEADASHCIQQILEAVLHCHQMGVVHRDLKPENLLLASKLK 138

Qy         196 KKLV-LENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTM 254
               | | :   :  : |  : :      | |: ||:|   |   ||  |:|: || |: :
Db         139 GAAVKLADFGLAIEVQGDQQAWFGFAGTPGYLSPEVLRKEA--YGKPVDIWACGVILYIL 196

Qy         255 LAGHYPFQDSEPVLLFGKIRRGAYALPA----GLSAPARCLVRCLLRREPAERLTATGIL 310
               | |: || | :   |: :|: |||  |:     ::  |: |:  :|   ||:|:||   |
Db         197 LVGYPPFWDEDQHKLYQQIKAGAYDFPSPEWDTVTPEAKNLINQMLTINPAKRITAHEAL 256

Qy         311 LHPWLRQ 317
                |||: |
Db         257 KHPWVCQ 263


RESULT 9
US-07-857-224B-25
; Sequence 25, Application US/07857224B
; Patent No. 5958784
;  GENERAL INFORMATION:
;    APPLICANT:  Benner, Steven A.
;    TITLE OF INVENTION:  Predicting Folded Structures of Proteins
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Steven A. Benner
;      STREET:  Hadlaubstrasse 151
;      CITY:  Zurich
;      STATE:  none
;      COUNTRY:  Switzerland
;      ZIP:  (note: this is an international post code) CH-8092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette, 1.4 Mb storage
;      COMPUTER:  Apple MacIntosh
;      OPERATING SYSTEM:  MacIntosh 7.0
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/857,224B
;      FILING DATE:  03/25/92
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:  none
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (International) 41 1 632 2830
;      TELEFAX:  (International) 41 1 262 2437
;      TELEX:  none
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  257
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Saccharomyces cerevisiae
;    FEATURE:  Protein kinase; Table 8 Column 28
;    PUBLICATION INFORMATION:
;      AUTHORS:
;      AUTHORS:  Hanks, S. K.
;      AUTHORS:  Quinn, A. M.
;      AUTHORS:  Hunter, T.
;      TITLE:  The protein kinase family
;      JOURNAL:  Science
;      VOLUME:  241
;      PAGES:  42-52
;      DATE:  1988

US-07-857-224B-25

  Query Match             13.9%;  Score 263;  DB 2;  Length 257;
  Best Local Similarity   28.8%;  Pred. No. 6.6e-19;
  Matches   72;  Conservative   43;  Mismatches  117;  Indels   18;  Gaps    6;

Qy          80 GRAYRALHCPTGTEYTCKVYPVQEALAVLE---------PYARLPPHKHVARPTEVLAGT 130
               |:   | |  || :   |:   :: ||  :           | ||  | |: :  :|:
Db          15 GKVKLAYHTTTGQKVALKIIN-KKVLAKSDMQGRIEREISYLRLLRHPHIIKLYDVIKSK 73

Qy         131 QLLYAFPTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVF 190
                  :          ::   :  | :: | ||   |:|: :|: :||:| :| ||||    :
Db          74 DEIIMVIEYAGNELFDYIVQRDKMSEQEARRFFQQIISAVEYCHRHKIVHRDLKPENLLL 133

Qy         191 ADRERKKLVLENLEDSCVLTGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLG 248
                  :    |:   |  | ::|   |  : | :|  | |  ||::| :  |:|   |||| |
Db         134 DEHLNVKIADFGL--SNIMT---DGNFLKTSCGSPNYAAPEVISGKL-YAGPEVDVWSCG 187

Qy         249 VALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATG 308
               | |: ||    || |    :||  |  | | ||  ||  |  |:: :|   |  |::
Db         188 VILYVMLCRRLPFDDESIPVLFKNISNGVYTLPKFLSPGAAGLIKRMLIVNPLNRISIHE 247

Qy         309 ILLHPWLRQD 318
               |:   | : |
Db         248 IMQDDWFKVD 257


RESULT 10
US-08-557-006C-40
; Sequence 40, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
;  APPLICANT:  Beri, Rajindar K.
;  APPLICANT:  Carling, David
;  APPLICANT:  Forder, Robert A.
;  TITLE OF INVENTION:  NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
;  FILE REFERENCE:  NGAP/PHM37588/UST
;  CURRENT APPLICATION NUMBER: US/08/557,006C
;  CURRENT FILING DATE:  1996-03-06
;  PRIOR APPLICATION NUMBER:  PCT/GB94/01093
;  PRIOR FILING DATE:  1994-05-20
;  PRIOR APPLICATION NUMBER:  GB 9310489.1
;  PRIOR FILING DATE:  1993-05-21
;  PRIOR APPLICATION NUMBER:  GB 9318010.7
;  PRIOR FILING DATE:  1993-08-31
;  NUMBER OF SEQ ID NOS:  44
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 40
;   LENGTH: 552
;   TYPE: PRT
;   ORGANISM: Rat
;   FEATURE:
;   NAME/KEY: gene
;   LOCATION: (1)..(1747)
;   OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
;   OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
;   OTHER INFORMATION: nucleotide 1765
US-08-557-006C-40

  Query Match             13.8%;  Score 260.5;  DB 3;  Length 552;
  Best Local Similarity   29.2%;  Pred. No. 3.6e-18;
  Matches   82;  Conservative   43;  Mismatches  119;  Indels   37;  Gaps    9;

Qy          67 RLGPYVL---LEPEEGGRAYRALHCPTGTEYTCKVYPVQE--ALAVLEPYAR------LP 115
               ::| |||   |     |:     |  || :   |:   |:  :| |:    |      |
Db          12 KIGHYVLGDTLGVGTFGKVKIGEHQLTGHKVAVKILNRQKIRSLDVVGKIKREIQNLKLF 71

Qy         116 PHKHVARPTEVLA-GTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCH 174
                 | |: :  :|:: |      :  |::   :    |: | ||  ||:|: :|: :||
Db          72 RHPHIIKLYQVISTPTDFFMVMEYVSGGELFDYICKHGRVEEVEARRLFQQILSAVDYCH 131

Qy         175 QHGLVLRDLKLCRFVFADRERKKLV---LEN-------LEDSCVLTGPDDSLWDKHACPA 224
               :| :| ||||    :   :   |:    | |       |  ||      .          |
```

```
Db        132 RHMVVHRDLKPENVLLDAQMNAKIADFGLSNMMSDGEFLRTSC------------GSPN 178

Qy        225 YVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGL 284

Db        179 YAAPEVISGRL-YAGPEVDIWSCGVILYALLCGTLPFDDEHVPTLFKKIRGGVFYIPEYL 237

Qy        285 SAPARCLVRCLLRREPAERLTATGILLHPWLRQD-PMPLAP 324

Db        238 NRSIATLLMHMLQVDPLKRATIKDIREHEWFKQDLPSYLFP 278


RESULT 11
US-08-688-988-29
; Sequence 29, Application US/08688988B
; Patent No. 6096545
; GENERAL INFORMATION:
;  APPLICANT: Lefebvre, Daniel D.
;  APPLICANT: Malboobi, Mohammad A.
;  TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
;  FILE REFERENCE: PPL96-03
;  CURRENT APPLICATION NUMBER: US/08/688,988B
;  CURRENT FILING DATE: 1996-07-31
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29
;   LENGTH: 354
;   TYPE: PRT
;   ORGANISM: Brassica napus
US-08-688-988-29

  Query Match             13.5%;  Score 256;  DB 3;  Length 354;
  Best Local Similarity   29.5%;  Pred. No. 5.5e-18;
  Matches   86;  Conservative   32;  Mismatches  104;  Indels   70;  Gaps   11;

Qy        117 HKHVARPTE-VLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQ 175

Db         57 HPNIIRFKEVVLTPTHLAIAMEYAAGGELFERICSAGRFSEDEARYFFQQLISGVSYCHA 116

Qy        176 HGLVLRDLKL---------------CRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKH 220

Db        117 MQICHRDLKLENTLLDGSPAPRLKICDFGYS----KSSLLHSRPKSTVGT---------- 162

Qy        221 ACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF----GKIRRG 276

Db        163 --PAYIAPEVL-SRREYDGKMADVWSCGVTLYVMLVGAYPFEDQEDPKNFRKTIQKIMAV 219

Qy        277 AYALP--AGLSAPARCLVRCLLRREPAERLTATGILLHPW-LRQDPMPLAPTRSHLW--- 330

Db        220 QYKIPDYVHISQDCKHLLSRIFVANSLKRITIAEIKKHPWFLKNLPRELTETAQAAYFKK 279

Qy        331 -----------EAAQVVPD------------GLGL----DEAREEEGDREVV 355

Db        280 ENPTFSPQTAEEIMKIVDDAKTPPPVSRSIGGFGWGGKGDEEEEEVDEEVV 331


RESULT 12
US-07-857-224B-18
; Sequence 18, Application US/07857224B
; Patent No. 5958784
;  GENERAL INFORMATION:
;    APPLICANT:  Benner, Steven A.
;    TITLE OF INVENTION:  Predicting Folded Structures of Proteins
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Steven A. Benner
;      STREET:  Hadlaubstrasse 151
;      CITY:  Zurich
;      STATE:  none
;      COUNTRY:  Switzerland
;      ZIP:  (note: this is an international post code) CH-8092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette, 1.4 Mb storage
;      COMPUTER:  Apple MacIntosh
;      OPERATING SYSTEM:  MacIntosh 7.0
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/857,224B
;      FILING DATE:  03/25/92
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:  none
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (International) 41 1 632 2830
;      TELEFAX:  (International) 41 1 262 2437
;      TELEX:  none
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  rat
;    FEATURE:  Protein kinase; Table 8 Column 20
;    PUBLICATION INFORMATION:
;      AUTHORS:
;      AUTHORS:  Hanks, S. K.
;      AUTHORS:  Quinn, A. M.
;      AUTHORS:  Hunter, T.
;      TITLE:  The protein kinase family
;      JOURNAL:  Science
;      VOLUME:  241
;      PAGES:  42-52
;      DATE:  1988
US-07-857-224B-18

  Query Match             13.4%;  Score 253;  DB 2;  Length 264;
  Best Local Similarity   29.4%;  Pred. No. 7.3e-18;
  Matches   77;  Conservative   49;  Mismatches  116;  Indels   20;  Gaps    8;

Qy         71 YVLLEPEEGGRAY----RALHCPTGTEYTCKVYPVQEALA----VLEPYA---RLPPHKH 119

Db          3 YQLFE-ELGKGAPSVVRRCVKVLAGQEYAAKIINTKKLSARDHQKLEREARICRLLKHPN 61

Qy        120 VARPTEVLAGTQLLYAFFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGL 178

Db         62 IVRLHDSISEEGHHYLIFDLVTGGELFEDIVAREYYSEADASHCIQQILEAVLHCHQMGV 121

Qy        179 VLRDLKLCRFVFADRERKKLV-LENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASY 237

Db        122 VHRDLKPENLLLASKLKGAAVKLADFGLAIEVEGEQQAWFGFAGTPGYLSPEVL--RKDP 179

Qy        238 SGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPA----GLSAPARCLVR 293

Db        180 YGKPVDLWACGVILYILLVGYPPFWDEDQHRLYQQIKAGAYDFPSPEWDTVTPEAKDLIN 239

Qy        294 CLLRREPAERLTATGILLHPWL 315

Db        240 KMLTINPSKRITAAEALKHPWI 261


RESULT 13
US-08-913-050A-7
; Sequence 7, Application US/08913050A
; Patent No. 5827726
;  GENERAL INFORMATION:
;    APPLICANT:  NEZU, Jun-ichi
;    TITLE OF INVENTION:  DNA ENCODING PROTEIN KINASE
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK, P.L.L.C.
;      STREET:  419 7th Street N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/913,050A
;      FILING DATE:  05-SEP-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 57104/1995
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP PCT/JP96/00660
;      FILING DATE:  15-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YUN, Allen C.
;      REGISTRATION NUMBER:  37,971
;      REFERENCE/DOCKET NUMBER:  NEZU=4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  433 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-913-050A-7

  Query Match             13.4%;  Score 252.5;  DB 2;  Length 433;
  Best Local Similarity   27.3%;  Pred. No. 1.7e-17;
  Matches   70;  Conservative   42;  Mismatches  129;  Indels   15;  Gaps    4;

Qy        117 HKHVARPTEVLAG--TQLLYAFFTRTHGDMHSLVRS--RHRIPEPEAAVLFRQMATALAH 172
              ||:| :  :||      | :|         |  :: |    || :|   | |:   | :
Db        107 HKNVIQLVDVLYNEEKQKMYMVMEYCVCGMQEMLDSVPEKRPPVCQAHGYFCQLIDGLEY 166

Qy        173 CHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILS 232
               |  |:| :|:|    :       |:    : ::       ||:       ||:  ||| :
Db        167 LHSQGIVHKDIKPGNLLLTTGGTLKISDLGVAEALHPFAADDTCRTSQGSPAFQPPEIAN 226

Qy        233 SRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLV 292
              ::||   |:|| || |: :  | |||:      ||  | :|:||:|    |   |:
Db        227 GLDTFSGFKVDIWSAGVTLYNITTGLYPFEGDNIYKLFENIGKGSYAIPGDCGPPLSDLL 286

Qy        293 RCLLRREPAERLTATGILLHPWLRQD--------PMPLAPTRSHLWEAAQVVP---DGLG 341
              : :|  |||:| :   |  | | |:        |:| :|     | :  |||   |  |
Db        287 KGMLEYEPAKRFSIRQIRQHSWFRKKHPPAEAPVPIPPSPDTKDRWRSMTVVPYLEDLHG 346

Qy        342 LDEAREEEGDREVVLY 357
               ||  :     : ::|
Db        347 ADEDEDLFDIEDDIIY 362


RESULT 14
US-08-749-902-5
; Sequence 5, Application US/08749902
; Patent No. 5985635
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    TITLE OF INVENTION:  NOVEL HUMAN SERINE/THREONINE
;    TITLE OF INVENTION:  PROTEIN KINASES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/749,902
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0150 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  433 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1480861
US-08-749-902-5

  Query Match             13.4%;  Score 252.5;  DB 2;  Length 433;
  Best Local Similarity   27.3%;  Pred. No. 1.7e-17;
  Matches   70;  Conservative   42;  Mismatches  129;  Indels   15;  Gaps    4;

Qy        117 HKHVARPTEVLAG--TQLLYAFFTRTHGDMHSLVRS--RHRIPEPEAAVLFRQMATALAH 172
              ||:| :  :||      | :|         |  :: |    || :|   | |:   | :
Db        107 HKNVIQLVDVLYNEEKQKMYMVMEYCVCGMQEMLDSVPEKRPPVCQAHGYFCQLIDGLEY 166

Qy        173 CHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILS 232
               |  |:| :|:|    :       |:    : ::       ||:       ||:  ||| :
Db        167 LHSQGIVHKDIKPGNLLLTTGGTLKISDLGVAEALHPFAADDTCRTSQGSPAFQPPEIAN 226

Qy        233 SRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLV 292
              ::||   |:|| || |: :  | |||:      ||  | :|:||:|    |   |:
Db        227 GLDTFSGFKVDIWSAGVTLYNITTGLYPFEGDNIYKLFENIGKGSYAIPGDCGPPLSDLL 286

Qy        293 RCLLRREPAERLTATGILLHPWLRQD--------PMPLAPTRSHLWEAAQVVP---DGLG 341
              : :|  |||:| :   |  | | |:        |:| :|     | :  |||   |  |
Db        287 KGMLEYEPAKRFSIRQIRQHSWFRKKHPPAEAPVPIPPSPDTKDRWRSMTVVPYLEDLHG 346

Qy        342 LDEAREEEGDREVVLY 357
               ||  :     : ::|
Db        347 ADEDEDLFDIEDDIIY 362


RESULT 15
US-08-688-988-32
; Sequence 32, Application US/08688988B
; Patent No. 6096545
; GENERAL INFORMATION:
;  APPLICANT: Lefebvre, Daniel D.
;  APPLICANT: Malboobi, Mohammad A.
;  TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
;  FILE REFERENCE: PPL96-03
;  CURRENT APPLICATION NUMBER: US/08/688,988B
;  CURRENT FILING DATE: 1996-07-31
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Brassica napus
US-08-688-988-32
```

```
  Query Match             13.1%;  Score 248;  DB 3;  Length 359;
  Best Local Similarity   28.6%;  Pred. No. 3.7e-17;
  Matches   84;  Conservative   33;  Mismatches  105;  Indels   72;  Gaps   11;

Qy        117 HKHVARPTE-VLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQ 175
              | :: |  | ||  | |  |       |::   :    |  | ||   |:|: : :::|| 
Db         57 HPNIIRFKEVVLTPTHLAIAMEYAAGGELFERICGAGRFSEDEARYFFQQLISGVSYCHA 116

Qy        176 HGLVLRDLKL---------------CRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKH 220
               :  |||||                | | ::    |  :| :   | | |          
Db        117 MQICHRDLKLENTLLDGSPAPRLKICDFGYS----KSSLLHSRPKSTVGT---------- 162

Qy        221 ACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF----GKIRRG 276
              |||: ||:| ||  | || |||||  || |: || | |||:| |    |     ||    
Db        163 --PAYIAPEVL-SRREYDGKMADVWSCGVTLYVMLVGAYPFEDQEDPKNFRKTIQKIMAV 219

Qy        277 AYALP--AGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD-PMPLAPTRSHLW--- 330
              | :|    :|   : |:  :      :|:|   |  |||  :: |  |  |    :    
Db        220 QYKIPDYVHISQDCKHLLSRIFVANSLKRITIAEIKKHPWFTKNLPRELTETAQAAYFKK 279

Qy        331 -----------EAAQVVPD-----------GLG------LDEAREEEGDREVV 355
                        |  ::| |           | |      |:   ||| | | |  
Db        280 ENPTFSAQTAEEIMKIVDDAKTPPPVSRSIGGFGWGGEGDLEGKEEEEVDEEEV 333


Search completed: January 15, 2004, 15:02:26
Job time : 22 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:11:12 ; Search time 7716.16 Seconds
(without alignments)
10916.435 Million cell updates/sec

Title: US-09-909-474D-1
Perfect score: 2059
Sequence: 1 gctctgagccccggcggcgc.....aaaaaaaaaaaaaaaaaaaa 2059

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		2055.8	99.8	2121	9	BC019363	BC019363 Homo sapi
2		2030.6	98.6	2283	9	BC027484	BC027484 Homo sapi
3		2003	97.3	2257	9	AK026945	AK026945 Homo sapi
4		1999.2	97.1	2116	6	AX099934	AX099934 Sequence
5		1990.8	96.7	2488	9	AY247738	AY247738 Homo sapi
6		1367	66.4	153170	9	HS1103G7	AL034548 Human DNA
7		1073.8	52.2	1077	6	AX166518	AX166518 Sequence
8		1054.6	51.2	1083	6	AX572896	AX572896 Sequence
9		1054.6	51.2	1083	9	AF250311	AF250311 Homo sapi
10		802.2	39.0	972	6	AX224734	AX224734 Sequence
11		661.4	32.1	1256	10	AF358868	AF358868 Mus muscu
12		645.2	31.3	1076	6	AX364906	AX364906 Sequence
13		645.2	31.3	1076	6	AX364921	AX364921 Sequence
14		626.4	30.4	2004	10	AB020967	AB020967 Rattus sp
15		546.8	26.6	1969	10	BC012955	BC012955 Mus muscu
16		533	25.9	541	6	AX381642	AX381642 Sequence
17	c	505.2	24.5	2559	6	AR274998	AR274998 Sequence
18	c	505.2	24.5	2559	9	HSU02019	U02019 Human AU-ri
19	c	505.2	24.5	2562	6	I24432	I24432 Sequence 4
20		499.2	24.2	75131	2	AC023022	AC023022 Homo sapi
21		330.2	16.0	396	6	AX245589	AX245589 Sequence
22	c	318	15.4	191503	2	AL928568	AL928568 Mus muscu
23	c	315.4	15.3	353	6	AX329832	AX329832 Sequence
24	c	297	14.4	75131	2	AC023022	AC023022 Homo sapi
25	c	295.6	14.4	235947	2	AC133773	AC133773 Rattus no
26	c	295.6	14.4	282849	2	AC132555	AC132555 Rattus no
27	c	278.2	13.5	331	11	G23310	G23310 human STS W
28		274.2	13.3	3319	6	AX572894	AX572894 Sequence
29		274.2	13.3	3319	9	AF250310	AF250310 Homo sapi
30		272.6	13.2	3317	9	AF205437	AF205437 Homo sapi
31		272.6	13.2	3338	9	HSM803696	AL832388 Homo sapi
32		261.4	12.7	2514	10	BC006800	BC006800 Mus muscu
33		258.2	12.5	1244	10	AF358866	AF358866 Mus muscu
34		254	12.3	1682	4	AY247741	AY247741 Bos tauru
35		253	12.3	2156	9	BC002637	BC002637 Homo sapi
36		253	12.3	3062	9	AY245544	AY245544 Homo sapi
37		253	12.3	4221	9	D87119	D87119 Homo sapien
38		250.6	12.2	1102	10	AF205438	AF205438 Rattus no
39		247	12.0	3324	6	AX281759	AX281759 Sequence
40		244.4	11.9	1032	4	CFC5FW	X99144 C.familiari
41		242.6	11.8	690	6	AX525590	AX525590 Sequence
42		241.2	11.7	1496	10	AF358867	AF358867 Mus muscu
43		241.2	11.7	3244	10	BC034338	BC034338 Mus muscu
44		234	11.4	675	9	HSC8FW	AJ000480 Homo sapi
45		222.2	10.8	2667	10	BC027159	BC027159 Mus muscu

ALIGNMENTS

RESULT 1
BC019363
LOCUS BC019363 2121 bp mRNA linear PRI 06-JUN-2003
DEFINITION Homo sapiens chromosome 20 open reading frame 97, mRNA (cDNA clone MGC:854 IMAGE:3528490), complete cds.
ACCESSION BC019363
VERSION BC019363.1 GI:17939414
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2121)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

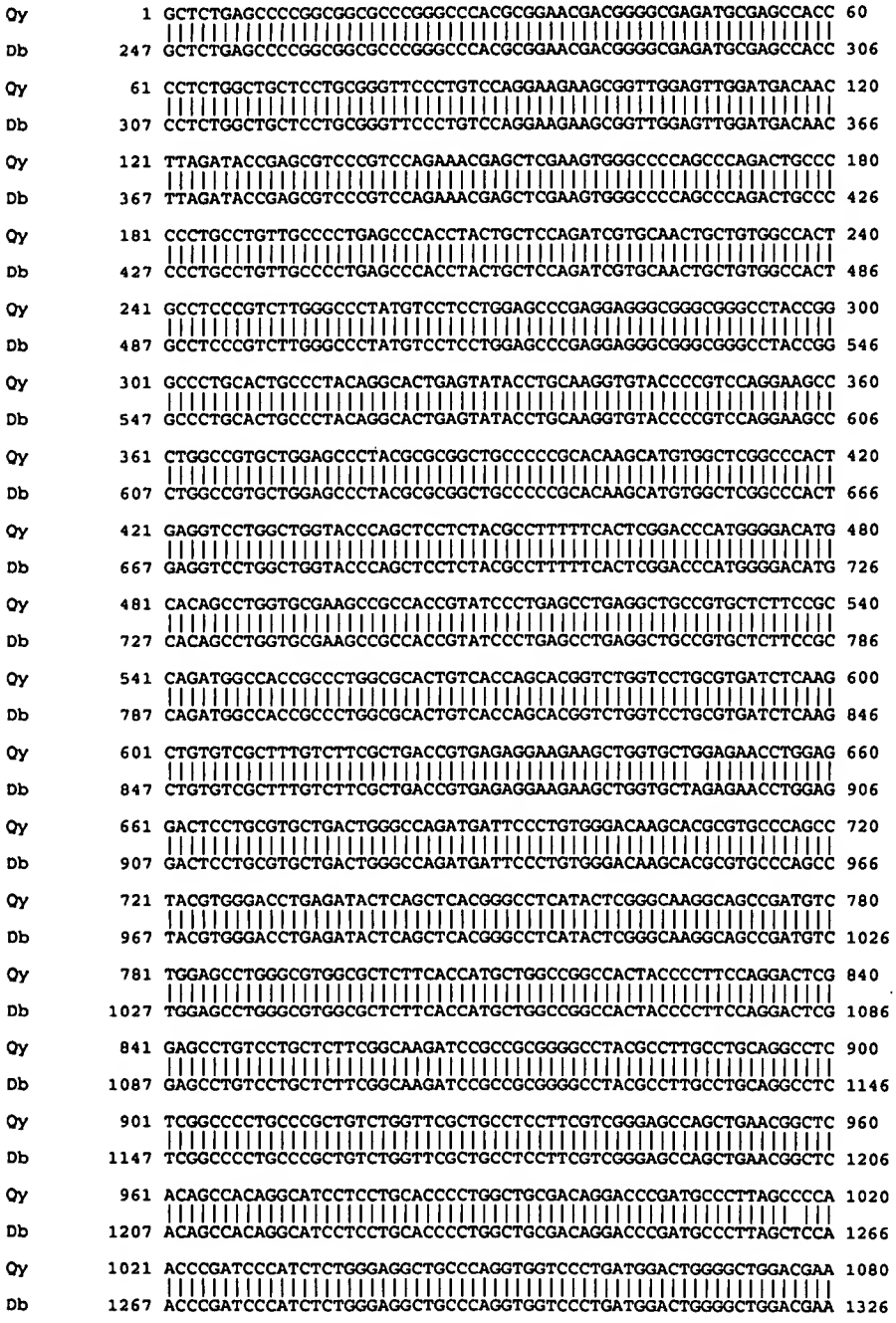
```
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2121)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 8 Row: g Column: 5
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 11056039.
FEATURES            Location/Qualifiers
     source         1. .2121
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="MGC:854 IMAGE:3528490"
                    /tissue_type="Muscle, rhabdomyosarcoma"
                    /clone_lib="NIH_MGC_17"
                    /lab_host="DH10B-R"
                    /note="Vector: pOTB7"
     gene           1. .2121
                    /gene="C20orf97"
                    /note="synonyms: dJ1103G7.3, SKIP3, TRB3"
                    /db_xref="LocusID:57761"
     CDS            108. .1184
                    /codon_start=1
                    /product="chromosome 20 open reading frame 97"
                    /protein_id="AAH19363.1"
                    /db_xref="GI:17939415"
                    /db_xref="LocusID:57761"
                    /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                    PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQ
                    EALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                    VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD
                    KHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG
                    AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ
                    VVPDGLGLDEARFEEGDREVVLYG"
BASE COUNT      459 a    645 c    584 g    433 t
ORIGIN

  Query Match             99.8%;  Score 2055.8;  DB 9;  Length 2121;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2057;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 119

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 179

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 239

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 299

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        300 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 359

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 419

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        420 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 479

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 539

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 599

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 659

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 719

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 779

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 839

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 899

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 959

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
```

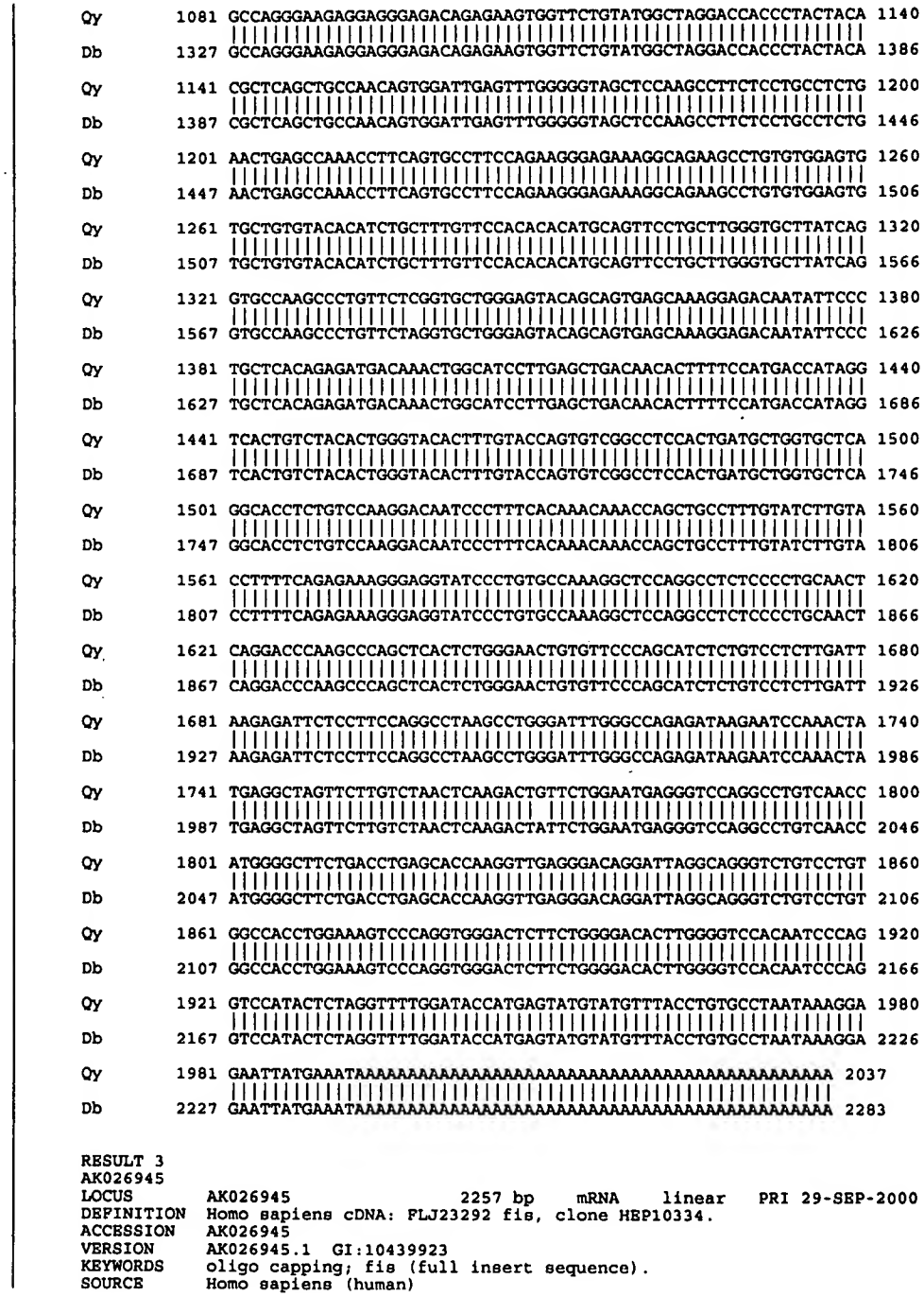
```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1019
Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1079
Qy       1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1139
Qy       1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1199
Qy       1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1259
Qy       1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1319
Qy       1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1379
Qy       1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1439
Qy       1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1499
Qy       1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1559
Qy       1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1619
Qy       1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1679
Qy       1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1739
Qy       1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1799
Qy       1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1859
Qy       1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1860 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1919
Qy       1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1979
Qy       1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2039
Qy       1981 GAATTATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2040 GAATTATGAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2099
Qy       2041 AAAAAAAAAAAAAAAAAA 2059
                ||||||||||||||||||
Db       2100 AAAAAAAAAAAAAAAAAA 2118
```

RESULT 2
BC027484

```
LOCUS       BC027484               2283 bp    mRNA    linear   PRI 08-APR-2002
DEFINITION  Homo sapiens, chromosome 20 open reading frame 97, clone MGC:34909
            IMAGE:5104452, mRNA, complete cds.
ACCESSION   BC027484
VERSION     BC027484.1  GI:20071610
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2283)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
            Richards, S., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 52 Row: f Column: 11.
FEATURES             Location/Qualifiers
     source          1. .2283
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="LocusID:57761"
                     /db_xref="taxon:9606"
                     /clone="MGC:34909 IMAGE:5104452"
                     /tissue_type="Cervix, carcinoma"
                     /clone_lib="NIH_MGC_12"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             295. .1371
                     /codon_start=1
                     /product="chromosome 20 open reading frame 97"
                     /protein_id="AAH27484.1"
                     /db_xref="GI:20071611"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQ
                     EALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                     VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD
                     KHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG
                     AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ
                     VVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      463 a    702 c    656 g    462 t
ORIGIN
```

```
  Query Match           98.6%;  Score 2030.6;  DB 9;  Length 2283;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 2033;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```





RESULT 3
AK026945
LOCUS AK026945 2257 bp mRNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens cDNA: FLJ23292 fis, clone HEP10334.
ACCESSION AK026945
VERSION AK026945.1 GI:10439923
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
            Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
            Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2257)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .2257
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEP10334"
                     /cell_line="HepG2"
                     /cell_type="hepatoma"
                     /clone_lib="HEP"
                     /note="cloning vector pME18SFL3"
     CDS             295. .1371
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB15597.1"
                     /db_xref="GI:10439924"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQ
                     EAPAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                     VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD
                     KHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG
                     AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ
                     VVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      435 a    705 c    659 g    458 t
ORIGIN

  Query Match           97.3%;  Score 2003;  DB 9;  Length 2257;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 2006;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 306

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CCTCTAGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 366

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 426

Qy        181 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 486

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        487 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 546

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 606

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              | ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        607 CCGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 666

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 726

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 786

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 846

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 906

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 966

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 1026

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 1086

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 1146

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1206

Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1266

Qy       1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1267 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGTCTGGGGCTGGACGAA 1326

Qy       1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1386

Qy       1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1446

Qy       1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1506

Qy       1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1566

Qy       1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1567 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1626
```

```
Qy       1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1686

Qy       1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1746

Qy       1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1806

Qy       1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1807 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1866

Qy       1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1926

Qy       1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1986

Qy       1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 2046

Qy       1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2047 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 2106

Qy       1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2107 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 2166

Qy       1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2226

Qy       1981 GAATTATGAAATAAAAAAAAAAAAAAAAAAAA 2011
             ||||||||||||||||||||||||||||||||
Db       2227 GAATTATGAAATAAAAAAAAAAAAAAAAAAAA 2257
```

```
RESULT 4
AX099934
LOCUS       AX099934               2116 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 16 from Patent WO0120004.
ACCESSION   AX099934
VERSION     AX099934.1  GI:13538944
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yue,H., Tang,Y.T., Bandman,O., Hillman,J.L., Baughn,M.R.,
            Azimzai,Y. and Lu,D.A.
  TITLE     Protein phosphatase and kinase proteins
  JOURNAL   Patent: WO 0120004-A 16 22-MAR-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2116
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID NO: 1271505CB1"
BASE COUNT      407 a    663 c    609 g    437 t
ORIGIN

  Query Match             97.1%;  Score 1999.2;  DB 6;  Length 2116;
  Best Local Similarity   99.9%;  Pred. No. 0;
```

```
  Matches 2001;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 172

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 232

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 292

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 352

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 412

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 472

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 532

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 592

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        593 CACAGCCTGGTGCGAACGCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 652

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 712

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 772

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 832

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 892

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 952

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 1012

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1072

Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1132

Qy       1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1192
```

```
Qy       1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1252

Qy       1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1312

Qy       1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1372

Qy       1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1432

Qy       1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1492

Qy       1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1552

Qy       1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1612

Qy       1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1672

Qy       1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1732

Qy       1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1792

Qy       1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1852

Qy       1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1912

Qy       1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1972

Qy       1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 2032

Qy       1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2092

Qy       1981 GAATTATGAAATAAAAAAAAAAAA 2004
             ||||||||||||||||||||||||
Db       2093 GAATTATGAAATAAAAAAAAAAA 2116
```

```
RESULT 5
AY247738
LOCUS       AY247738               2488 bp    mRNA    linear   PRI 20-APR-2003
DEFINITION  Homo sapiens TRB3 protein mRNA, complete cds.
ACCESSION   AY247738
VERSION     AY247738.1  GI:30025661
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2488)
  AUTHORS   Shan,Y.X. and Yu,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-2003) School of Life Science, Institute of
            Genetics, Fudan University, Handan Road 220, Shanghai 200433, PR
            China
FEATURES             Location/Qualifiers
     source          1. .2488
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             541. .1617
                     /codon_start=1
                     /product="TRB3 protein"
                     /protein_id="AAP04407.1"
                     /db_xref="GI:30025662"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQ
                     EALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                     VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD
                     KHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG
                     AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ
                     VVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      478 a    768 c    746 g    496 t
ORIGIN
```

Query Match 96.7%; Score 1990.8; DB 9; Length 2488;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 552

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 612

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 672

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 732

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        733 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 792

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 852

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        853 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 912

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        913 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 972

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 1032

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 1092
```

```
Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
Db       1093 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 1152

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
Db       1153 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 1212

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
Db       1213 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 1272

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
Db       1273 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 1332

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
Db       1333 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 1392

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
Db       1393 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1452

Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
Db       1453 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1512

Qy       1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
Db       1513 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1572

Qy       1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
Db       1573 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1632

Qy       1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
Db       1633 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1692

Qy       1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
Db       1693 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1752

Qy       1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
Db       1753 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1812

Qy       1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
Db       1813 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1872

Qy       1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
Db       1873 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1932

Qy       1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
Db       1933 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1992

Qy       1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
Db       1993 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 2052

Qy       1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
Db       2053 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 2112

Qy       1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
Db       2113 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 2172

Qy       1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
Db       2173 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 2232

Qy       1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
Db       2233 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 2292

Qy       1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
Db       2293 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 2352

Qy       1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
Db       2353 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 2412

Qy       1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
Db       2413 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2472

Qy       1981 GAATTATGAAATAA 1994
Db       2473 GAATTATGAAATAA 2486
```

RESULT 6
HS1103G7

LOCUS	HS1103G7 153170 bp DNA linear PRI 08-FEB-2001
DEFINITION	Human DNA sequence from clone RP5-1103G7 on chromosome 20p12.2-13. Contains up to three novel genes, the gene for a novel protein similar to mouse VMP, the gene for a novel protein kinase domains containing protein similar to phosphoprotein C8FW and rat NIPK, and the SOX22 gene for SRY (sex-determining region Y)-box 22. Contains five CpG islands, ESTs, STSs and GSSs, complete sequence.
ACCESSION	AL034548
VERSION	AL034548.25 GI:7263904
KEYWORDS	HTG; CpG island; NIPK; protein kinase; SOX22; SRY; VMP.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 153170)
AUTHORS	Blakey,S.
TITLE	Direct Submission
JOURNAL	Submitted (05-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT	On Mar 19, 2000 this sequence version replaced gi:5541861. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence is the entire insert of clone RP5-1103G7 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr20 RP5-1103G7 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm

```
          VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1. .153170
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="p12.2-13"
                     /clone="RP5-1103G7"
                     /clone_lib="RPCI-5"
     repeat_region   1. .61
                     /note="Alu repeat: matches 2. .62 of consensus"
     repeat_region   79. .129
                     /note="L1MA9 repeat: matches 6221. .6271 of consensus"
     repeat_region   130. .173
                     /note="22 copies 2 mer ag 79% conserved"
     repeat_region   175. .873
                     /note="L1MA9 repeat: matches 5519. .6217 of consensus"
     repeat_region   902. .1109
                     /note="L1MC5 repeat: matches 7715. .7921 of consensus"
     repeat_region   1110. .1480
                     /note="THE1C repeat: matches 1. .371 of consensus"
     misc_feature    1502. .1733
                     /note="match: STS: Em:Z66755"
     repeat_region   1734. .1888
                     /note="MLT1H repeat: matches 387. .547 of consensus"
     repeat_region   1919. .2101
                     /note="MLT1H repeat: matches 115. .298 of consensus"
     misc_feature    complement(join(2146. .2243,126791. .126834))
                     /note="match: GSS: Em:B59651"
     repeat_region   2154. .2193
                     /note="20 copies 2 mer aa 85% conserved"
     misc_feature    complement(join(2194. .2243,126792. .126871))
                     /note="match: GSS: Em:AQ377604"
     repeat_region   2699. .2917
                     /note="L1M4 repeat: matches 2875. .3081 of consensus"
     repeat_region   2938. .3062
                     /note="L1M4 repeat: matches 3120. .3249 of consensus"
     repeat_region   3136. .3430
                     /note="AluSg repeat: matches 1. .295 of consensus"
     repeat_region   3482. .3777
                     /note="AluSq repeat: matches 1. .296 of consensus"
     repeat_region   3811. .4126
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   4310. .4608
                     /note="L1MD2 repeat: matches 5861. .6167 of consensus"
     repeat_region   4609. .5026
                     /note="MSTA repeat: matches 1. .424 of consensus"
     repeat_region   5027. .5200
                     /note="L1MD2 repeat: matches 5684. .5861 of consensus"
     repeat_region   5201. .5501
                     /note="AluSx repeat: matches 1. .307 of consensus"
     repeat_region   5502. .5727
                     /note="L1MD2 repeat: matches 5451. .5684 of consensus"
     repeat_region   5798. .5865
                     /note="34 copies 2 mer aa 66% conserved"
     repeat_region   6056. .6643
                     /note="MER67C repeat: matches 122. .710 of consensus"
     repeat_region   6648. .6707
                     /note="30 copies 2 mer ca 88% conserved"
     repeat_region   6710. .6824
                     /note="MER50 repeat: matches 10. .134 of consensus"
     repeat_region   6825. .7124
                     /note="AluSx repeat: matches 7. .306 of consensus"
     repeat_region   7126. .7415
                     /note="L1MA4A repeat: matches 5619. .5919 of consensus"
     repeat_region   7419. .7569
                     /note="AluSg/x repeat: matches 148. .298 of consensus"
     misc_feature    7576. .7889
                     /note="match: GSS: Em:AQ180303"
     misc_feature    complement(7590. .7894)
                     /note="match: GSS: Em:AQ701808"
     misc_feature    7862. .8140
                     /note="match: GSS: Em:AQ123077"
     misc_feature    8019. .8317
                     /note="match: GSS: Em:AQ746749"
     misc_feature    8019. .8151
                     /note="match: GSS: Em:AQ617736"
     repeat_region   8152. .8316
                     /note="MER4B repeat: matches 416. .574 of consensus"
     repeat_region   8349. .8643
                     /note="MER4D repeat: matches 362. .689 of consensus"
     repeat_region   8644. .8942
                     /note="AluY repeat: matches 1. .299 of consensus"
     repeat_region   8943. .9113
                     /note="MER4D repeat: matches 194. .362 of consensus"
     repeat_region   9113. .9201
                     /note="MLT2FA repeat: matches 377. .450 of consensus"
     repeat_region   9226. .9311
                     /note="L1MB4 repeat: matches 5775. .5862 of consensus"
     repeat_region   9312. .9651
                     /note="MER1B repeat: matches 1. .337 of consensus"
     repeat_region   9653. .9750
                     /note="L1MA5 repeat: matches 5856. .5947 of consensus"
     repeat_region   9751. .10047
                     /note="AluSx repeat: matches 1. .298 of consensus"
     repeat_region   10048. .10369
                     /note="L1MA5 repeat: matches 5947. .6300 of consensus"
     repeat_region   10488. .10529
                     /note="21 copies 2 mer at 88% conserved"
     repeat_region   10532. .10653
                     /note="L1MC4 repeat: matches 7673. .7800 of consensus"
     repeat_region   10687. .10752
                     /note="33 copies 2 mer at 68% conserved"
     repeat_region   11681. .11742
                     /note="31 copies 2 mer ta 71% conserved"
     misc_feature    13519. .13972
                     /note="match: GSS: Em:AQ224309"
     repeat_region   14208. .14554
                     /note="LTR16B repeat: matches 110. .462 of consensus"
     repeat_region   15554. .15872
                     /note="L1ME1 repeat: matches 5615. .5910 of consensus"
     repeat_region   15873. .16174
                     /note="AluSp repeat: matches 16. .313 of consensus"
     repeat_region   16175. .16382
                     /note="L1ME1 repeat: matches 5400. .5615 of consensus"
     repeat_region   16383. .16423
                     /note="L1ME repeat: matches 5893. .5933 of consensus"
     repeat_region   16595. .16894
                     /note="AluY repeat: matches 1. .302 of consensus"
     repeat_region   16895. .17021
                     /note="L1MA10 repeat: matches 6196. .6318 of consensus"
     repeat_region   17155. .17587
                     /note="L1ME repeat: matches 5289. .5733 of consensus"
     repeat_region   17587. .18157
                     /note="L1M1 repeat: matches -1389. .-790 of consensus"
     repeat_region   18158. .18195
                     /note="19 copies 2 mer tg 97% conserved"
     repeat_region   18197. .18230
                     /note="AluSg/x repeat: matches 165. .197 of consensus"
     repeat_region   18231. .18823
                     /note="MER52A repeat: matches 1. .672 of consensus"
     misc_feature    complement(18827. .19180)
                     /note="match: GSS: Em:AQ599398"
     repeat_region   19689. .19998
                     /note="AluSg repeat: matches 1. .309 of consensus"
     repeat_region   20287. .20543
                     /note="AluSx repeat: matches 2. .307 of consensus"
     repeat_region   20552. .20609
                     /note="MLT1J repeat: matches 255. .318 of consensus"
     repeat_region   20592. .20756
                     /note="MLT1E repeat: matches 256. .421 of consensus"
     repeat_region   20895. .21012
                     /note="L1PBa repeat: matches -1128. .-1008 of consensus"
     repeat_region   21009. .21168
                     /note="L1PBa repeat: matches -1537. .-1377 of consensus"
```

```
     repeat_region   21192. .21277
                     /note="L1MA7 repeat: matches 5864. .5949 of consensus"
     repeat_region   21280. .21627
                     /note="L2 repeat: matches 2250. .2617 of consensus"
     misc_feature    21628. .21884
                     /note="match: GSS: Em:B41180"
     misc_feature    21628. .21744
                     /note="match: GSS: Em:AQ058160"

  Query Match             66.4%;  Score 1367;  DB 9;  Length 153170;
  Best Local Similarity   98.9%;  Pred. No. 1.8e-252;
  Matches 1376;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy        604 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 663
              || | |||    ||| |     ||   ||||||||||||||||||||||||||||||||
Db     150847 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 150906

Qy        664 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     150907 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 150966

Qy        724 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     150967 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 151026

Qy        784 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151027 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 151086

Qy        844 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151087 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 151146

Qy        904 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151147 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 151206

Qy        964 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151207 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 151266

Qy       1024 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151267 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 151326

Qy       1084 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151327 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 151386

Qy       1144 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAAC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151387 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAAC 151446

Qy       1204 TGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGC 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151447 TGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGC 151506

Qy       1264 TGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAGGTG 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151507 TGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAGGTG 151566

Qy       1324 CCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151567 CCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGC 151626

Qy       1384 TCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCA 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151627 TCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCA 151686

Qy       1444 CTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGC 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151687 CTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGC 151746

Qy       1504 ACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151747 ACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCT 151806

Qy       1564 TTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAG 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151807 TTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAG 151866

Qy       1624 GACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAG 1683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151867 GACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAG 151926

Qy       1684 AGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGA 1743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151927 AGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGA 151986

Qy       1744 GGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATG 1803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151987 GGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATG 152046

Qy       1804 GGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGC 1863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152047 GGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGC 152106

Qy       1864 CACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTC 1923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152107 CACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTC 152166

Qy       1924 CATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAA 1983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152167 CATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAA 152226

Qy       1984 TTATGAAATAA 1994
              ||||||||||||
Db     152227 TTATGAAATAA 152237


RESULT 7
AX166518
LOCUS       AX166518               1077 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 9 from Patent WO0138503.
ACCESSION   AX166518
VERSION     AX166518.1  GI:14546863
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
            Planagan,P. and Clary,D.S.
  TITLE     Novel human protein kinases and protein kinase-like enzymes
  JOURNAL   Patent: WO 0138503-A 9 31-MAY-2001;
            Sugen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1077
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      174 a    373 c    325 g    205 t
ORIGIN

  Query Match             52.2%;  Score 1073.8;  DB 6;  Length 1077;
  Best Local Similarity   99.8%;  Pred. No. 1.9e-196;
  Matches 1075;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy        109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
```

```
Qy        169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 360

Qy        409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Qy        649 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Qy        709 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Qy        769 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Qy        829 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840

Qy        889 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900

Qy        949 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960

Qy       1009 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020

Qy       1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 1077
```

```
RESULT 8
AX572896
LOCUS       AX572896               1083 bp    DNA     linear   PAT 29-NOV-2002
DEFINITION  Sequence 3 from Patent WO02053743.
ACCESSION   AX572896
VERSION     AX572896.1  GI:26004968
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dower,S. and Quanstrom,E.
  TITLE     Mammalian tribbles signaling pathways and methods and reagents
            related thereto
  JOURNAL   Patent: WO 02053743-A 3 11-JUL-2002;
            Interleukin Genetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1083
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      176 a    374 c    327 g    206 t
ORIGIN
```

Query Match 51.2%; Score 1054.6; DB 6; Length 1083;
Best Local Similarity 99.1%; Pred. No. 9.6e-193;
Matches 1073; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

```
Qy         49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy        109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
              |||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGGTGCCCCCGCACAAGCATGTG 360

Qy        409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGC------TGACCGTGAGAGGAAGAAGCTG 642
              |||||||||||||||||||||||||||||||      |||||||||| ||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGACCGTGAGAAGAAGAAGCTG 600

Qy        643 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 660

Qy        703 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 720

Qy        763 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 780
```

```
Qy        823 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGCGGGGCCTAC 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGCGGGGCCTAC 840

Qy        883 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 900

Qy        943 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 960

Qy       1003 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1020

Qy       1063 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1080

Qy       1123 TAG 1125
              |||
Db       1081 TAG 1083
```

```
RESULT 9
AF250311
LOCUS       AF250311                1083 bp    mRNA    linear   PRI 01-JUN-2001
DEFINITION  Homo sapiens SKIP3 mRNA, complete cds.
ACCESSION   AF250311
VERSION     AF250311.1  GI:14276268
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1083)
  AUTHORS   Kiss-Toth,E., Wyllie,D.H., Qwarnstrom,E.E. and Dower,S.K.
  TITLE     Identification of pro-inflammatory cytokine signalling network
            components by transcription expression screening
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1083)
  AUTHORS   Kiss-Toth,E., Wyllie,D.H., Qwarnstrom,E.E. and Dower,S.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAR-2000) Division of Molecular & Genetic Medicine,
            University of Sheffield, Royal Hallamshire Hospital, Floor M,
            Sheffield S10-2JF, UK
FEATURES             Location/Qualifiers
     source          1. .1083
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             1. .1083
                     /codon_start=1
                     /product="SKIP3"
                     /protein_id="AAK58175.1"
                     /db_xref="GI:14276269"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQ
                     EALAVLEPYARVPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                     VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRDREKKKLVLENLEDSCVLTGPDDSL
                     WDKHACPAYVGPEILSSRASYSGKAADVWSLGVALPTMLAGHYPPQDSEPVLLPGKIR
                     RGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEA
                     AQVVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      176 a    374 c    327 g    206 t
ORIGIN

  Query Match           51.2%;  Score 1054.6;  DB 9;  Length 1083;
  Best Local Similarity   99.1%;  Pred. No. 9.6e-193;
  Matches 1073;  Conservative    0;  Mismatches    4;  Indels    6;  Gaps    1;

Qy         49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy        109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGGTGCCCCCGCACAAGCATGTG 360

Qy        409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGC------TGACCGTGAGAGGAAGAAGCTG 642
              |||||||||||||||||||||||||||||||      |||||||||| ||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGACCGTGAGAAGAAGAAGCTG 600

Qy        643 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 660

Qy        703 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 720

Qy        763 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 780

Qy        823 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGCGGGGCCTAC 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGCGGGGCCTAC 840

Qy        883 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 900

Qy        943 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 960

Qy       1003 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1020

Qy       1063 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1080

Qy       1123 TAG 1125
              |||
```

```
Db       1081 TAG 1083


RESULT 10
AX224734
LOCUS       AX224734                 972 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 21 from Patent WO0160991.
ACCESSION   AX224734
VERSION     AX224734.1  GI:15554836
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Buford,N., Gandhi,A.R., Patterson,C., Khan,F.A., Yue,H.,
            Hafalia,A., Shih,L.L., Tribouley,C.M., Yao,M.G., Burrill,J.D.,
            Marcus,G.A., Zingler,K.A., Lu,D.A., Bandman,O., Policky,J.L.,
            Griffin,J.A., Thornton,M., Nguyen,D.B. and Walsh,R.T.
  TITLE     Human kinases
  JOURNAL   Patent: WO 0160991-A 21 23-AUG-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .972
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472328CB1"
BASE COUNT      167 a    329 c    294 g    182 t
ORIGIN

  Query Match             39.0%;  Score 802.2;  DB 6;  Length 972;
  Best Local Similarity   87.7%;  Pred. No. 3.6e-144;
  Matches  944;  Conservative    0;  Mismatches   28;  Indels  105;  Gaps    2;

Qy          49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
               ||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db           1 ATGCGAGCCACCCCTTTGGCTGCTTCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy         109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy         169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy         229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy         289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
               |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy         349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
               ||||||||||||||||||||||||| ||||| |||||||||||||| ||||||||||||||
Db         301 GTCCAGGAAGCCCTGGCCGTGCTGGAACCCTATGCGCGGCTGCCCCCGGACAAGCATGTG 360

Qy         409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
               ||| |||||||||||| |||||||||||||| || ||||||||||||||||| |||||||
Db         361 GCTTGGCCCACTGAGGGCCTGGCTGGTACCCAACTTCTCTACGCCTTTTTCACTTGGACC 420

Qy         469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
               |||||||||||||||| ||| |  |
Db         421 CATGGGGACATGCACAGGCTGATAGG---------------------------------- 446

Qy         529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
                           |  |||  ||   ||||||||| |||||
Db         447 ----------------GCACACGCCCTGTGCGCACTGTGACCAG--------------- 474

Qy         589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 648
                                                   ||||||||||||||||||
Db         475 ------------------------------------ACACGGAAGAAGCTGGTGCTG 495

Qy         649 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 708
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         496 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 555

Qy         709 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 768
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         556 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 615

Qy         769 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 828
               |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         616 GCAGCCAATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 675

Qy         829 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 888
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         676 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 735

Qy         889 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 948
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         736 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 795

Qy         949 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1008
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         796 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 855

Qy        1009 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1068
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         856 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 915

Qy        1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 1125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         916 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 972


RESULT 11
AF358868
LOCUS       AF358868                 1256 bp    mRNA    linear   ROD 01-JUN-2002
DEFINITION  Mus musculus TRB-3 mRNA, complete cds.
ACCESSION   AF358868
VERSION     AF358868.1  GI:21304713
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1256)
  AUTHORS   Kiss-Toth,E., Dempsey,C., Jozsa,V., Caunt,J., Oxley,K.M.,
            Bagstaff,S.M., Wyllie,D.H., Harte,M., O'Neill,L.A.J.,
            Qwarnstrom,E.E. and Dower,S.K.
  TITLE     Mammalian homologs of Drosophila tribbles (htrb) control mitogen
            activated protein kinase signaling
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1256)
  AUTHORS   Kiss-Toth,E., Dempsey,C., Jozsa,V., Caunt,J., Oxley,K.M.,
            Bagstaff,S.M., Wyllie,D.H., Harte,M., O'Neill,L.A.J.,
            Qwarnstrom,E.E. and Dower,S.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-2001) Division of Genomic Medicine, University of
            Sheffield, Royal Hallamshire Hospital, Floor M, Glossop Road,
            Sheffield S10-2JF, UK
FEATURES             Location/Qualifiers
     source          1. .1256
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS             182. .1246
                     /note="mTRB-3"
                     /codon_start=1
                     /product="TRB-3"
                     /protein_id="AAM45476.1"
                     /db_xref="GI:21304714"
                     /translation="MRATPLAASADVSCRKKPLEFDDNIDAKCPVLKRVRDEPEPGPL
```

```
                     PSLLPPSPPPASDLSPAVAPATRLGPYILLEREQGSCSYRALHCPTGTEYTCKVYPAS
                     EAQAVLAPYARLPTHQHVARPTEVLLGSRLLYIPFTKTHGDLHSLVRSRRGIPESEAA
                     GLFRQMASAVAHCHKHGLVLRDLKLRRFVFSNCERTKLVLENLEDACVMTGSDDSLWD
                     THACPAYVGPEILSSRPSYSGKAADVWSLGVALFTMLAGRYPFHDSEPVLLFGKIRXG
                     TFALPEGLSAPARCLIRCLLRKEPSERLVALGILLHPWLREDHGRVSPPQSDRREMDQ
                     VVPDGPQLEEAEEGEVGLYG"
BASE COUNT      231 a    390 c    385 g    249 t      1 others
ORIGIN

  Query Match           32.1%;  Score 661.4;  DB 10;  Length 1256;
  Best Local Similarity   75.7%;  Pred. No. 4.8e-117;
  Matches  818;  Conservative    0;  Mismatches  262;  Indels    0;  Gaps    0;

Qy         17 GCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACCCCTCTGGCTGCTCCTG 76
Db        150 GGGCGCAGGCGGCTCTGAGGCTCCAGGACAAGATGCGAGCTACACCTCTGGCTGCTTCTG 209

Qy         77 CGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTC 136
Db        210 CTGATGTTTCCTGCAGGAAGAAACCGTTGGAGTTTGATGACAATATTGATGCCAAGTGTC 269

Qy        137 CCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCC 196
Db        270 CAGTCCTAAAACGAGTGAGAGATGAGCCTGAGCCCGGACCACTCCCCAGCCTGCTGCCCC 329

Qy        197 TGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGC 256
Db        330 CCAGCCCACCTCCCGCCTCAGACTTGTCACCTGCTGTGGCCCCTGCAACTCGACTGGGGC 389

Qy        257 CCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTA 316
Db        390 CTTATATCCTTTTGGAACGAGAGCAAGGCAGCTGCAGCTATCGAGCCCTGCACTGCCCCA 449

Qy        317 CAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGC 376
Db        450 CAGGCACAGAGTACACCTGCAAGGTGTACCCTGCCAGCGAGGCCCAGGCGGTGCTGGCAC 509

Qy        377 CCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTA 436
Db        510 CTTATGCCCGGCTGCCTACCCACCAGCATGTGGCCCGTCCCACAGAGGTCCTGCTGGGCT 569

Qy        437 CCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAA 496
Db        570 CTCGGCTCCTTTACATCTTTTTCACGAAGACCCATGGGGACTTGCACAGCCTGGTGCGCA 629

Qy        497 GCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCC 556
Db        630 GCCGCCGCGGTATCCCAGAGTCCGAGGCTGCCGGGCTCTTCCGGCAGATGGCTAGTGCCG 689

Qy        557 TGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCT 616
Db        690 TGGCACACTGCCACAAGCACGGGCTTGTCTTGCGCGACCTCAAGCTGCGTCGCTTTGTCT 749

Qy        617 TCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGA 676
Db        750 TCAGCAACTGTGAGAGGACGAAGCTGGTGCTGGAGAACCTGGAAGATGCCTGCGTGATGA 809

Qy        677 CTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGA 736
Db        810 CTGGATCAGATGACTCTCTGTGGGACACGCATGCGTGCCCTGCCTACGTGGGACCAGAGA 869

Qy        737 TACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGG 796
Db        870 TACTCAGCTCCCGGCCATCCTACTCTGGCAAAGCGGCTGATGTCTGGAGCCTGGGTGTGG 929

Qy        797 CGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCT 856
Db        930 CGCTCTTCACCATGCTGGCTGGCAGATACCCATTCCACGACTCTGAGCCAGTCCTGCTCT 989

Qy        857 TCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCT 916
Db        990 TTGGCAAGATCCGTANAGGGACCTTTGCCCTGCCTGAGGGCCTATCAGCCCCAGCCCGCT 1049

Qy        917 GTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCC 976
Db       1050 GTCTGATCCGCTGTCTCCTCCGCAAGGAACCTTCAGAGCGACTTGTGGCCCTGGGAATCC 1109

Qy        977 TCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCTCT 1036
Db       1110 TCTTGCATCCCTGGTTGAGAGAGGATCACGGCCGAGTCTCTCCTCCACAGTCTGACCGAA 1169

Qy       1037 GGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGG 1096
Db       1170 GGGAGATGGACCAGGTGGTCCCAGATGGGCCACAGCTGGAGGAGGCTGAGGAAGGGGAGG 1229


RESULT 12
AX364906
LOCUS       AX364906               1076 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 57 from Patent WO0206315.
ACCESSION   AX364906
VERSION     AX364906.1  GI:18696795
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mintz,L., Freilich,S. and Bernstein,J.
  TITLE     Novel nucleic acid and amino acid sequences
  JOURNAL   Patent: WO 0206315-A 57 24-JAN-2002;
            Compugen Ltd. (IL)
FEATURES             Location/Qualifiers
     source          1. .1076
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      172 a    352 c    346 g    194 t     12 others
ORIGIN

  Query Match           31.3%;  Score 645.2;  DB 6;  Length 1076;
  Best Local Similarity   98.2%;  Pred. No. 6.2e-114;
  Matches  670;  Conservative    3;  Mismatches    6;  Indels    3;  Gaps    2;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
Db        281 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 340

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
Db        341 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 400

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
Db        401 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 460

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
Db        461 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 520

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
Db        521 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 580

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
Db        581 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 640

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
Db        641 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 700

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
Db        701 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 760
```

```
Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||::||||||||||||||||||||||||||||||||||||||||
Db        761 CACAGCCTGGTGCGAAGCSSCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 820

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 880

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              |||||||||||||||||||||||||||||||| |||||||||||||||||  | |
Db        881 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG-AGAAGCTGGTGCTGGAGA--CCTGGA 937

Qy        661 GACTCCTGCGTGCTGACTGGGC 682
              |||||:||||||||||||||||
Db        938 GACTCSTGCGTGCTGACTGGGC 959


RESULT 13
AX364921
LOCUS       AX364921               1076 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 72 from Patent WO0206315.
ACCESSION   AX364921
VERSION     AX364921.1  GI:18696810
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mintz,L., Freilich,S. and Bernstein,J.
  TITLE     Novel nucleic acid and amino acid sequences
  JOURNAL   Patent: WO 0206315-A 72 24-JAN-2002;
            Compugen Ltd. (IL)
FEATURES             Location/Qualifiers
     source          1. .1076
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      172 a    352 c    346 g    194 t     12 others
ORIGIN

  Query Match             31.3%;  Score 645.2;  DB 6;  Length 1076;
  Best Local Similarity   98.2%;  Pred. No. 6.2e-114;
  Matches  670;  Conservative    3;  Mismatches    6;  Indels    3;  Gaps    2;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 340

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 400

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 460

Qy        181 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 520

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        521 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 580

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 640

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        641 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 700

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 760

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||::||||||||||||||||||||||||||||||||||||||||
Db        761 CACAGCCTGGTGCGAAGCSSCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 820

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 880

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              |||||||||||||||||||||||||||||||| |||||||||||||||||  | |
Db        881 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG-AGAAGCTGGTGCTGGAGA--CCTGGA 937

Qy        661 GACTCCTGCGTGCTGACTGGGC 682
              |||||:||||||||||||||||
Db        938 GACTCSTGCGTGCTGACTGGGC 959


RESULT 14
AB020967
LOCUS       AB020967               2004 bp    mRNA    linear   ROD 14-MAY-1999
DEFINITION  Rattus sp. mRNA for kinase, complete cds.
ACCESSION   AB020967
VERSION     AB020967.1  GI:4827158
KEYWORDS    kinase; NIPK.
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Mayumi-Matsuda,K., Kojima,S., Suzuki,H. and Sakata,T.
  TITLE     Identification of a novel kinase-like gene induced during neuronal
            cell death
  JOURNAL   Biochem. Biophys. Res. Commun. 258 (2), 260-264 (1999)
  MEDLINE   99262087
   PUBMED   10329375
REFERENCE   2  (bases 1 to 2004)
  AUTHORS   Matsuda,K.M., Kojima,S. and Nakayama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1998) Keiko Mayumi Matsuda, SHIONOGI and Co.
            Ltd., Shionogi Institute for Medical Science; 2-5-1 Mishima,
            Settsu-shi, Osaka 566-0022, Japan
            (E-mail:keiko.matsuda@shionogi.co.jp, Tel:81-6-6382-2612(ex.465),
            Fax:81-6-6382-2598)
FEATURES             Location/Qualifiers
     source          1. .2004
                     /organism="Rattus sp."
                     /mol_type="mRNA"
                     /db_xref="taxon:10118"
                     /cell_line="PC12"
                     /cell_type="pheochromocytoma"
     gene            1. .2004
                     /gene="NIPK"
     CDS             154. .1203
                     /gene="NIPK"
                     /note="neuronal cell death inducible putative kinase
                     (NIPK);induced by NGF-depletion"
                     /codon_start=1
                     /product="kinase"
                     /protein_id="BAA77582.1"
                     /db_xref="GI:4827159"
                     /translation="MRATSLAASADVPCRKKPLEFDDNIDVECPVLKRVRDEPEPGPT
                     PSLPPASDLSPAVAPATRLGPYILLEREQGNCTYRALHCPTGTEYTCKVYPASEAQAV
                     LAPYARLPTHQHVARPTEVLLGSQLLYTFFTKTHGDLHSLVRSRRGIPEPEAAALFRQ
                     MASAVAHCHKHGLILRDLKLRRFVFSNCERTKLVLENLEDACVMTGPDDSLWDKHACP
                     AYVGPEILSSRPSYSGRAADVWSLGVALFTMLAGRYPFQDSEPALLFGKIRRGTFALP
                     EGLSASARCLIRCLLRREPSERLVALGILLHPWLREDCSQVSPPRSDRREMDQVVPDG
                     PQLEEAEEGEVGLYG"
     polyA_site      2004
```

```
                     /gene="NIPK"
                     /note="27 a nucleotides"
BASE COUNT      410 a    563 c    590 g    441 t
ORIGIN

  Query Match           30.4%;  Score 626.4;  DB 10;  Length 2004;
  Best Local Similarity   72.1%;  Pred. No. 2.9e-110;
  Matches  883;  Conservative    0;  Mismatches  311;  Indels   30;  Gaps    4;

Qy         17 GCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACCCCTCTGGCTGCTCCTG 76
              | || | |||       |  ||  | || | |||||||||||||  |||||||||||  |||
Db        122 GGGCGCAGGCGGCTCTGAGACCTCAGGACAAGATGCGAGCCACATCTCTGGCTGCTTCTG 181

Qy         77 CGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTC 136
              | | |    |  | |||||||||| | |||||||||| |||||||||  | |||  |||| |||
Db        182 CCGATGTTCCCTGCAGGAAGAAACCGTTGGAGTTTGATGACAACATTGATGTCGAGTGTC 241

Qy        137 CCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCC 196
              | |||    |||||||    ||  || |||   ||||   |||   |||||
Db        242 CAGTCTTAAAACGAGTGAGAGATGAGCCTGAGCCTGGACCAACCCCC---------- 288

Qy        197 TGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGC 256
                |||| |||| | ||  ||||     || ||||||||||| |||  |||| || ||||
Db        289 --AGCCTACCTCCCGCCTCAGACTTGTCACCTGCTGTGGCCCCTGCAACCCGGCTGGGGC 346

Qy        257 CCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTA 316
              ||||| ||||  |||| |  ||| | |||     |  |||| ||||||||||||||||| |
Db        347 CCTATATCCTTTTGGAACGAGAGCAAGGCAACTGCACCTATCGGGCCCTGCACTGCCCCA 406

Qy        317 CAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGC 376
              ||||||| ||||| |||||||||||||| |||| |   || |||| ||| |||||||  |
Db        407 CAGGCACAGAGTACACCTGCAAGGTGTATCCCGCCAGCGAGGCCCAGGCGGTGCTGGCAC 466

Qy        377 CCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTA 436
              | || ||  | |||||  | ||| |||| ||||| || ||||| |||||||||   ||
Db        467 CGTATGCAAGACTGCCTACCCACCAGCACGTGGCCCGTCCCACAGAGGTCCTGCTGGGCT 526

Qy        437 CCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAA 496
              | |||||||| ||| |||||||||   |||||||||||| ||||||| |||||||| |
Db        527 CTCAGCTCCTATACACCTTTTTCACGAAGACCCATGGGGACTTGCACAGTCTGGTGCGCA 586

Qy        497 GCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCC 556
              |||||| | |||||||| ||||| |||||||||| ||| ||||| |||||||| |  ||
Db        587 GCCGCCGCGGTATCCCCGAGCCCGAGGCTGCCGCGCTTTTCCGACAGATGGCTAGTGCGG 646

Qy        557 TGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCT 616
              |||| ||||| ||| ||||||| ||   || |||| || |||||| || |||| |||||||
Db        647 TGGCACACTGCCACAAGCACGGGCTCATCTTGCGCGACCTCAAGTTGCGTCGATTTGTCT 706

Qy        617 TCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGA 676
              ||    || |||||||||| |||| |||||||||||||||||||||  |||| ||| |||
Db        707 TCAGCAACTGTGAGAGGACGAAGTTGGTGCTGGAGAACCTGGAGGATGCCTGTGTGATGA 766

Qy        677 CTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGA 736
              |||| ||||| || || |||||||||||||||||||||||| ||||| ||||||||  ||||
Db        767 CTGGACCAGACGACTCTCTGTGGGACAAGCACGCGTGCCCTGCCTATGTGGGACCAGAGA 826

Qy        737 TACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGG 796
              |||||||||| ||| | || ||||| |||   || || ||||||||||||||||||||||
Db        827 TACTCAGCTCCCGGCCATCCTACTCCGGCAGAGCGGCTGATGTCTGGAGCCTGGGCGTGG 886

Qy        797 CGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCT 856
              |||||||||||||||||| |||| ||||| |||||||||| || || | ||||||||
Db        887 CGCTCTTCACCATGCTGGCTGGCCGATACCCGTTCCAGGACTCTGAACCAGCCCTGCTCT 946

Qy        857 TCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCT 916
              | |||||||||||  | ||| |||  ||| ||||||  ||||| || ||| |||| || |
Db        947 TTGGCAAGATCCGTAGAGGGACCTTTGCCCTGCCTGAGGGCCTATCAGCCTCTGCTCGAT 1006

Qy        917 GTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCC 976
              | ||  | ||||| ||||| ||  |||||||  | || || ||    |||   || ||||
Db       1007 GCCTCATCCGCTGTCTCCTCCGCAGGGAGCCTTCAGAGCGACTTGTGGCCCTGGGAATCC 1066

Qy        977 TCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCTCT 1036
              | || || |||||| || |  ||||     |  |   |  | ||  | || ||  | |
Db       1067 TGCTACATCCCTGGTTGAGGGAGGATTGCAGCCAAGTCTCTCCTCCACGGTCTGACCGAA 1126

Qy       1037 GGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGG 1096
              |||||   | |||||| ||||| ||||| |    ||||||| || ||   ||||| ||||
Db       1127 GGGAGATGGACCAGGTTGTCCCAGATGGGCCACAGCTGGAGGAGGCAGAGGAAGGGGAG- 1185

Qy       1097 GAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCT-ACTACACGCTCAGCTGCCAAC 1155
                        ||||  |||||||||||||  ||||| | |||   |  |||||| | |
Db       1186 ----------GTGGGGCTGTATGGCTAGAGCCACCTTAACTGGCCCTTCAGCTCCAAGG 1234

Qy       1156 AGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGCCAAACCT 1215
              |||  ||||||   || ||  |  |   |      |||   | ||  |  ||| ||||
Db       1235 TGTGAGTTGAGTACTGGTGTCTCAGCTTTCGCCTGCCT--TTTTGGGCCAAGCTAAACTG 1292

Qy       1216 TCAGTGCCTTCCAGAAGGGAGAAA 1239
              | |||||||| | | ||| |||||
Db       1293 TAAGTGCCTTTCTGGAGGAAGAAA 1316


RESULT 15
BC012955
LOCUS       BC012955                1969 bp    mRNA    linear   ROD 16-APR-2003
DEFINITION  Mus musculus induced in fatty liver dystrophy 2, mRNA (cDNA clone
            MGC:18731 IMAGE:3980838), complete cds.
ACCESSION   BC012955
VERSION     BC012955.1  GI:15277944
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1969)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1969)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
```

```
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 23 Row: k Column: 9
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES              Location/Qualifiers
     source           1. .1969
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="FVB/N"
                      /db_xref="taxon:10090"
                      /clone="MGC:18731 IMAGE:3980838"
                      /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                      model. 10 month old virgin mouse. Taken by biopsy."
                      /clone_lib="NCI_CGAP_Mam1"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     gene             1. .1969
                      /gene="Ifld2"
                      /note="synonym: TRB-3"
                      /db_xref="LocusID:228775"
                      /db_xref="MGI:1345675"
     CDS              177. .1163
                      /codon_start=1
                      /product="TRB-3"
                      /protein_id="AAH12955.1"
                      /db_xref="GI:15277945"
                      /db_xref="LocusID:228775"
                      /translation="MRATPLAASADVSCRKKPLEFDDNIDAKCPVLKRVRDEPEPGPL
                      PSLLPPSPPPASDLSPAVAPATRLGPYILLEREQGSCSYRALHCPTGTEYTCKVYPAS
                      EAQAVLAPYARLPTHQHVARPTEVLLGSRLLYIPFTKTHGDLHSLVRSRRGIPEPEAA
                      GLFRQMASAVAHCHKHGLVLRDLKLRRFVFSNCERTKLVLENLEDACVMTGSDDSLWD
                      KHACPAYVGPEILSSRPSYSEPVLLFGKIRRGTFALPEGLSAPARCLIRCLLRKEPSE
                      RLVALGILLHPWLREDHGRVSPPQSDRREMDQVVPDGPQLEEAEEGEVGLYG"
BASE COUNT      413 a    560 c    572 g    424 t
ORIGIN

  Query Match           26.6%;  Score 546.8;  DB 10;  Length 1969;
  Best Local Similarity   61.9%;  Pred. No. 6.1e-95;
  Matches 1247;  Conservative    0;  Mismatches  572;  Indels  197;  Gaps   16;

Qy         17 GCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACCCCTCTGGCTGCTCCTG 76
Db        145 GGGCGCAGGCGGCTCTGAGGCTCCAGGACAAGATGCGAGCTACACCTCTGGCTGCTTCTG 204

Qy         77 CGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTC 136
Db        205 CTGATGTTTCCTGCAGGAAGAAACCGTTGGAGTTTGATGACAATATTGATGCCAAGTGTC 264

Qy        137 CCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCC 196
Db        265 CAGTCCTAAAACGAGTGAGAGATGAGCCTGAGCCCGGACCACTCCCCAGCCTGCTGCCCC 324

Qy        197 TGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGC 256
Db        325 CCAGCCCACCTCCCGCCTCAGACTTGTCACCTGCTGTGGCCCCTGCAACTCGACTGGGGC 384

Qy        257 CCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTA 316
Db        385 CTTATATCCTTTTGGAACGAGAGCAAGGCAGCTGCAGCTATCGAGCCCTGCACTGCCCCA 444

Qy        317 CAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGC 376
Db        445 CAGGCACAGAGTACACCTGCAAGGTGTACCCTGCCAGCGAGGCCCAGGCGGTGCTGGCAC 504

Qy        377 CCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTA 436

Db        505 CTTATGCCCGGCTGCCTACCCACCAGCATGTGGCCCGTCCCACAGAGGTCCTGCTGGGCT 564

Qy        437 CCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAA 496
Db        565 CTCGGCTCCTTTACATCTTTTTCACGAAGACCCATGGGGACTTGCACAGCCTGGTGCGCA 624

Qy        497 GCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCC 556
Db        625 GCCGCCGCGGTATCCCAGAGCCCGAGGCTGCCGGGCTCTTCCGGCAGATGGCTAGTGCCG 684

Qy        557 TGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCT 616
Db        685 TGGCACACTGCCACAAGCACGGGCTTGTCTTGCGCGACCTCAAGCTGCGTCGCTTTGTCT 744

Qy        617 TCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGA 676
Db        745 TCAGCAACTGTGAGAGGACGAAGCTGGTGCTGGAGAACCTGGAAGATGCCTGCGTGATGA 804

Qy        677 CTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGA 736
Db        805 CTGGATCAGATGACTCTCTGTGGGACAAGCATGCGTGCCCTGCCTACGTGGGACCAGAGA 864

Qy        737 TACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGG 796
Db        865 TACTCAGCTCCCGGCCATCCT---------------------------------------- 885

Qy        797 CGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCT 856
Db        886 ----------------------------------------ACTCTGAGCCAGTCCTGCTCT 906

Qy        857 TCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCT 916
Db        907 TTGGCAAGATCCGTAGAGGGACCTTTGCCCTGCCTGAGGGCCTATCAGCCCCAGCCCGCT 966

Qy        917 GTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCC 976
Db        967 GTCTGATCCGCTGTCTCCTCCGCAAGGAACCTTCAGAGCGACTTGTGGCCCTGGGAATCC 1026

Qy        977 TCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCTCT 1036
Db       1027 TCTTGCATCCCTGGTTGAGAGAGGATCACGGCCGAGTCTCTCCTCCACAGTCTGACCGAA 1086

Qy       1037 GGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGG 1096
Db       1087 GGGAGATGGACCAGGTGGTCCCAGATGGGCCACAGCTGGAGGAGGCTGAGGAAGGGGAG- 1145

Qy       1097 GAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAGCTGCCAACA 1156
Db       1146 -----------GTGGGGCTGTACGGCTAGGGCCACCTTACTGGCCCCTCAGCTCCAAGGT 1194

Qy       1157 GTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGCCAAACCTT 1216
Db       1195 GTGAGTTGAGTACCTGAGTCTCAGCTTTCTCCTGACTT--TTTGGGCCAAGCTAAACCTT 1252

Qy       1217 CAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGTACACATCT 1276
Db       1253 AAGTGCCTTTCTGGAGGAAGAA----------------------------ACAGCCGG 1282

Qy       1277 GCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGCCCTGTTC 1336
Db       1283 CGTGCCTTATTCGTTCCTGTGCCTAGTGGGGTGATGCTCCCTGAGTGCCACGCCCTGCTC 1342

Qy       1337 TCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGCTCACAGAGATGAC 1396
Db       1343 TAGGTGCTGTGAACAGCAAAGGAAAAAGAGGGAGAAATATACCTTGCTGGCCAAGTTGCC 1402

Qy       1397 AAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCACTGTCTACACTGG 1456
Db       1403 ACAGTGCCG----------------------CATGCTCTCCGGCCAGCCCCTGCCTTGG 1439

Qy       1457 GTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCAAG 1516
```

```
Db       1440 GCACGTTTCCTACC-GGGGCTGTCTTCTCTGGTGCTGGT--------------TCATGG 1483

Qy       1517 GACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAGAGAAAGG 1576

Db       1484 GACAGTCCCTTTTATCAGTGCCCCAGCTGCC-TTCTGTCGTGTGCCTTTAAA---CAAAG 1539

Qy       1577 GAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTC---CCCTGCAACTCAGGACCCAAGCC 1633

Db       1540 GAAACATCTCTGTGCCAAAGACCAAAGTCCTCACTCACCATGTGCCTCAGGACCACAAAT 1599

Qy       1634 CAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCT 1693

Db       1600 CAGGCCGGGCTGGGATTGTGGCCTCCAGGCAGGAGAACCTGTTTGAAGA---------- 1648

Qy       1694 TCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAGTTCT 1753

Db       1649 ---AGACTTGGCTGTGGGATTCAAGCCAAAAGGAAGAGACTCCAATAAAGAGGCATTT-- 1703

Qy       1754 TGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGG-GGCTTCTG 1812

Db       1704 ----GACTCAAGGAATAGTCTGGAATAAGGATCCGGGTCTGTTTACCCTGGTAGCCTCCG 1759

Qy       1813 ACCTGAGCACCAAGGTTG-AGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGA 1871

Db       1760 GCTGAAGCACCAAAGTTGGAGGGGGAGGACCAGGCAGGGTCTGTCCTACTGCCTACAG-- 1817

Qy       1872 AAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACTCT 1931

Db       1818 ----ACAATCTCAGGCCCTGTTTTGGGGACTTTTAGTACCCAGGCCCACAGTCAGTTTGA 1873

Qy       1932 AGGTTTTGGATACCATGAGTATGTATGTTTA-CCTGTGCCTAATAAAGGAGAATTATGAA 1990

Db       1874 TGACTTTGGTCTGGTTAAGTGTGTATTTTCATCATGTGCCCAATAAAAGAGAGCTGTGTG 1933

Qy       1991 ATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2026

Db       1934 GTTTGATATGTGAAAAAAAAAAAAAAAAAAAAAAAA 1969
```

Search completed: January 16, 2004, 14:49:00
Job time : 7734.16 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:55:46 ; Search time 41 Seconds
(without alignments)
2253.240 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 1891
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	756	40.0	343	11	Q8K017	Q8k017 mus musculu
2	754.5	39.9	343	6	Q28283	Q28283 canis famil
3	753.5	39.8	343	4	Q92519	Q92519 homo sapien
4	753	39.8	343	11	Q8K4K3	Q8k4k3 mus musculu
5	749.5	39.6	364	11	Q9EQL6	Q9eql6 rattus norv
6	744.5	39.4	372	4	Q96RU8	Q96ru8 homo sapien
7	742.5	39.3	372	4	Q9H2Y8	Q9h2y8 homo sapien
8	742.5	39.3	372	11	Q91W04	Q91w04 mus musculu
9	729.5	38.6	372	11	Q8K4K4	Q8k4k4 mus musculu
10	653	34.5	257	11	Q8R2V8	Q8r2v8 mus musculu
11	611.5	32.3	224	4	O15180	O15180 homo sapien
12	435	23.0	484	5	Q9V3Z1	Q9v3z1 drosophila
13	326	17.2	218	11	Q8BFS7	Q8bfs7 mus musculu
14	326	17.2	379	11	Q8BJR9	Q8bjr9 mus musculu
15	326	17.2	388	11	Q8BZX3	Q8bzx3 mus musculu
16	315	16.7	461	10	Q9LGV5	Q9lgv5 oryza sativ
17	314.5	16.6	706	3	Q8J2N0	Q8j2n0 fusarium ox
18	306	16.2	474	10	Q94DT7	Q94dt7 oryza sativ
19	302.5	16.0	646	4	Q96CV1	Q96cv1 homo sapien
20	299	15.8	914	5	Q19469	Q19469 caenorhabdi
21	297.5	15.7	446	10	Q8H5S1	Q8h5s1 oryza sativ
22	297.5	15.7	778	4	Q8TDC2	Q8tdc2 homo sapien
23	297	15.7	504	11	Q8K0J7	Q8k0j7 mus musculu
24	297	15.7	880	3	Q9Y880	Q9y880 cochliobolu
25	296	15.7	671	3	Q96WI7	Q96wi7 trichoderma
26	295	15.6	794	4	Q8TDC3	Q8tdc3 homo sapien
27	292.5	15.5	651	13	Q91821	Q91821 xenopus lae
28	291	15.4	602	3	P87209	P87209 kluyveromyc
29	289.5	15.3	698	5	O61298	O61298 halocynthia
30	289	15.3	688	5	Q95UF4	Q95uf4 ancylostoma
31	288	15.2	477	10	Q9LWM6	Q9lwm6 oryza sativ
32	285	15.1	514	10	Q9M726	Q9m726 lycopersico
33	284	15.0	504	10	P93113	P93113 cucumis sat
34	284	15.0	512	10	P92958	P92958 arabidopsis
35	283.5	15.0	643	11	Q61804	Q61804 mus musculu
36	283.5	15.0	643	11	Q61846	Q61846 mus musculu
37	283.5	15.0	651	4	Q14680	Q14680 homo sapien
38	283	15.0	438	10	Q8LK24	Q8lk24 glycine max
39	283	15.0	512	10	P92968	P92968 arabidopsis
40	282	14.9	444	10	Q9FUK2	Q9fuk2 arabidopsis
41	281.5	14.9	419	4	Q9H7H6	Q9h7h6 homo sapien
42	281	14.9	535	10	Q8RWD2	Q8rwd2 arabidopsis
43	280	14.8	339	10	Q39868	Q39868 glycine max
44	280	14.8	444	10	Q8RWC9	Q8rwc9 arabidopsis
45	280	14.8	480	10	Q9LUP6	Q9lup6 arabidopsis

ALIGNMENTS

RESULT 1
Q8K017
ID Q8K017 PRELIMINARY; PRT; 343 AA.
AC Q8K017;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TRB-2 (C5FW ORF protein homolog).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Cerebellum, and Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; BC034338; AAH34338.1; -.
DR EMBL; AK044747; BAC32063.1; -.
DR EMBL; AK080064; BAC37820.1; -.
DR EMBL; AK082329; BAC38467.1; -.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 2.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.

```
SQ   SEQUENCE   343 AA;  38772 MW;  9418B7AC19FCC23F CRC64;

  Query Match             40.0%;  Score 756;  DB 11;  Length 343;
  Best Local Similarity   46.9%;  Pred. No. 1.8e-55;
  Matches  164;  Conservative   48;  Mismatches  108;  Indels   30;  Gaps    6;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRATA 61
              |:||:        || |  : :: | : |  :       |     | |   ||| |: :
Db          5 RSTPITIARYGRSRNKTQDFEE-LSSIRSAE------PSQSFSPNLGSPSPPETPNLSHC 57

Qy         62 VATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPV---QEALAVLEPYARLPPHK 118
                |   | :| |:|||| ||    :||:|  :| |  |||: :   ||:||   |   | |
Db         58 V---SCIGKYLLLEPLEGDHVFRAVHLHSGEELVCKVPEISCYQESLA---PCFCLSAHS 111

Qy        119 HVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGL 178
              :: : ||:| |     | || |::|||| ||:  :: | ||| || |:|:|:|||   ||
Db        112 NINQITEILLGETKAYVFFERSYGDMHSFVRTCKKLREEEAARLFYQIASAVAHCHDGGL 171

Qy        179 VLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYS 238
              ||||||| :|:| | || :: ||:|||: :| | |||| ||| |||| ||||::  |||
Db        172 VLRDLKLRKFIFKDEERTRVKLESLEDAYILRGDDDSLSDKHGCPAYVSPEILNTSGSYS 231

Qy        239 GKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRR 298
              |||||||||||| |:||| | ||| | ||  || ||||| : :|  ||  |:||:| :|||
Db        232 GKAADVWSLGVMLYTMLVGRYPFHDIEPSSLFSKIRRGQFNIPETLSPKAKCLIRSILRR 291

Qy        299 EPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREE 348
              ||:||||:  || |||   |                |   | |  || ::
Db        292 EPSERLTSQEILDHPWFSTD-------------FSVSNSGFGAKEACDQ 327


RESULT 2
Q28283
ID   Q28283      PRELIMINARY;      PRT;   343 AA.
AC   Q28283;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   C5FW ORF protein.
GN   C5FW ORF.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RX   MEDLINE=97067069; PubMed=8910471;
RA   Wilkin F., Savonet V., Radulescu A., Petermans J., Dumont J.E.,
RA   Maenhaut C.;
RT   "Identification and Characterization of Novel Genes Modulated in the
RT   Thyroid of Dogs Treated with Methimazole and Propylthiouracil.";
RL   J. Biol. Chem. 271:28451-28457(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RX   MEDLINE=98000262; PubMed=9342215;
RA   Wilkin F., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,
RA   Dumont J.E., Maenhaut C.;
RT   "Characterization of a phosphoprotein whose mRNA is regulated by the
RT   mitogenic pathways in dog thyroid cells.";
RL   Eur. J. Biochem. 248:660-669(1997).
DR   EMBL; X99144; CAA67581.1; -.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   343 AA;  38786 MW;  BF8D1300DACB84FA CRC64;

  Query Match             39.9%;  Score 754.5;  DB 6;  Length 343;
  Best Local Similarity   47.3%;  Pred. No. 2.4e-55;
  Matches  167;  Conservative   50;  Mismatches  115;  Indels   21;  Gaps    6;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRATA 61
              |:||:        || |  : :: | : |  :       |     | |   ||| |: :
Db          5 RSTPITIARYGRSRNKTQDFEE-LSSIRSAE------PSQSFSPNLGSPSPPETPNLSHC 57

Qy         62 VATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPV---QEALAVLEPYARLPPHK 118
                |   | :| |:|||| ||    :||:|  :| |  |||: :   ||:||   |   | |
Db         58 V---SCIGKYLLLEPLEGDHVFRAVHLHSGEELVCKVPDISCYQESLA---PCFCLSAHS 111

Qy        119 HVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGL 178
              :: : ||:: |     | || |::|||| ||:  :: | ||| || |:|:|:|||   ||
Db        112 NINQITEIILGETKAYVFFERSYGDMHSFVRTCKKLREEEAARLFYQIASAVAHCHDGGL 171

Qy        179 VLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYS 238
              ||||||| :|:| | || :: ||:|||: :| | |||| ||| |||| ||||::  |||
Db        172 VLRDLKLRKFIFKDEERTRVKLESLEDAYILRGDDDSLSDKHGCPAYVSPEILNTSGSYS 231

Qy        239 GKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRR 298
              |||||||||||| |:||| | ||| | ||  || ||||| : :|  ||  |:||:| :|||
Db        232 GKAADVWSLGVMLYTMLVGRYPFHDIEPSSLFSKIRRGQFNIPETLSPKAKCLIRSILRR 291

Qy        299 EPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGD 351
              ||:||||:  || |||   |      : |:     |: | | |  ||  |
Db        292 EPSERLTSQEILDHPWFSTD-----FSVSNSGYGAKEVSDQLVPDVNMEENLD 339


RESULT 3
Q92519
ID   Q92519      PRELIMINARY;      PRT;   343 AA.
AC   Q92519;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   GS3955 (GS3955 protein).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cancellous bone;
RA   Ohno I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Uterus;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D87119; BAA13250.1; -.
DR   EMBL; BC002637; AAH02637.1; -.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   343 AA;  38800 MW;  BF8B7366DACB84FA CRC64;

  Query Match             39.8%;  Score 753.5;  DB 4;  Length 343;
  Best Local Similarity   47.3%;  Pred. No. 2.9e-55;
  Matches  167;  Conservative   50;  Mismatches  115;  Indels   21;  Gaps    6;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRATA 61
              |:||:        || |  : :: | : |  :       |     | |   ||| |: :
Db          5 RSTPITIARYGRSRNKTQDFEE-LSSIRSAE------PSQSFSPNLGSPSPPETPNLSHC 57

Qy         62 VATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPV---QEALAVLEPYARLPPHK 118
                |   | :| |:|||| ||    :||:|  :| |  |||: :   ||:||   |   | |
Db         58 V---SCIGKYLLLEPLEGDHVFRAVHLHSGEELVCKVPDISCYQESLA---PCFCLSAHS 111

Qy        119 HVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGL 178
              :: : ||:: |     | || |::|||| ||:  :: | ||| || |:|:|:|||   ||
```

```
Db        112 NINQITEIILGETKAYVFFERSYGDMHSFVRTCKKLREEEAARLFYQIASAVAHCHDGGL 171

Qy        179 VLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYS 238
              ||||||| :|:| | || :: ||:|||: :| | |||| ||| ||||| ||||:: |||
Db        172 VLRDLKLRKPIFKDEERTRVKLESLEDAYILRGDDDSLSDKHGCPAYVSPEILNTSGSYS 231

Qy        239 GKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRR 298
              |||||||||||| |:||| | ||| | ||  || ||||| : :|  ||  |:||:| :|||
Db        232 GKAADVWSLGVMLYTMLVGRYPFHDIEPSSLFSKIRRGQFNIPETLSPKAKCLIRSILRR 291

Qy        299 EPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGD 351
              ||:||||:  || |||   |      : |:    |: | | |  |   ||  |
Db        292 EPSERLTSQEILDHPWFSTD-----FSVSNSAYGAKEVSDQLVPDVNMEENLD 339


RESULT 4
Q8K4K3
ID   Q8K4K3      PRELIMINARY;      PRT;   343 AA.
AC   Q8K4K3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   TRB-2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kiss-Toth E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA   Bagstaff S.M., Wyllie D.H., Harte M., O'Neill L.A.J., Qwarnstrom E.E.,
RA   Dower S.K.;
RT   "Mammalian homologs of Drosophila tribbles (htrb) control mitogen
RT   activated protein kinase signaling.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF358867; AAM45477.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   343 AA;  38758 MW;  0B3965B8B2087D74 CRC64;

  Query Match             39.8%;  Score 753;  DB 11;  Length 343;
  Best Local Similarity   46.6%;  Pred. No. 3.1e-55;
  Matches  163;  Conservative   49;  Mismatches  108;  Indels   30;  Gaps    6;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRATA 61
              |:||:       || |  : :: | : |  :       |      | |   |||  |: :
Db          5 RSTPITIARYGRSRNKTQDPEE-LSSIRSAE------PSQSPSPNLGSPSPPETPNLSHC 57

Qy         62 VATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPV---QEALAVLEPYARLPPHK 118
              |   | :| |:|||| ||   :||:|  :| |  |||: :   ||:||    |   | |
Db         58 V---SCIGKYLLLEPLEGDHVFRAVHLHSGEELVCKVFEISCYQESLA---PCFCLSAHS 111

Qy        119 HVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGL 178
              :: : ||:| |    | || |::||||| ||:  :: | ||| || |:|:|:|||| ||
Db        112 NINQITEILLGETKAYVFFERSYGDMHSFVRTCKKLREEEAARLFYQIASAVAHCHDGGL 171

Qy        179 VLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYS 238
              ||||||| :|:| | || :: ||:|||: :| | |||| ||| ||||| ||||:: |||
Db        172 VLRDLKLRKFIFKDEERTRVKLESLEDAYILRGDDDSLSDKHGCPAYVSPEILNTSGSYS 231

Qy        239 GKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRR 298
              |||||||||||| ::||| | ||| | ||  || ||||| : :|  ||  |:||:| :|||
Db        232 GKAADVWSLGVMVYTMLVGRYPFHDIEPSSLFSKIRRGQFNIPETLSPKAKCLIRSILRR 291

Qy        299 EPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREE 348
              ||:||||:  || |||   |              |  | |  || ::
Db        292 EPSERLTSQEILDHPWFSTD-------------FSVSNSGFGAKEACDQ 327


RESULT 5
Q9EQL6
ID   Q9EQL6      PRELIMINARY;      PRT;   364 AA.
AC   Q9EQL6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   G-protein-coupled receptor induced protein GIG2 (Fragment).
GN   GIG2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RA   Mayhaus M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,
RA   Nitsch R.M.;
RT   "Identification of a novel nuclear factor Gig2, as an m1-acetylcholine
RT   receptor-induced gene.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF205438; AAG35664.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Receptor; Transferase.
FT   NON_TER       1      1
SQ   SEQUENCE   364 AA;  40377 MW;  8F9167FB76DFCD37 CRC64;

  Query Match             39.6%;  Score 749.5;  DB 11;  Length 364;
  Best Local Similarity   46.2%;  Pred. No. 6.6e-55;
  Matches  163;  Conservative   46;  Mismatches  119;  Indels   25;  Gaps    5;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLP---LSPPTAP-- 56
              |   |  ||   :  |||     |||:     |:    |||  |  |   |||| :|
Db         10 RGPALLFPAARGTPAKRL-----LDTDDAAAVAAKC---PRLSECSNPPDYLSPPGSPCS 61

Qy         57 -----------DRATAVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
                          :    ||:  |:||   |     |||   || |  |||:|::
Db         62 PQPPPSAQGTGGSCVSAPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKVFPIKHYQ 121

Qy        106 AVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQ 165
              : || :|| |:::    ||: |    | || :  ||||| |||| |: | ||| ||:|
Db        122 DKIRPYTQLPSHRNITGIVEVILGESKAYVFFEKDFGDMHSYVRSRKRLREEEAARLFKQ 181

Qy        166 MATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAY 225
              : :|:|||||  :|| |||| :|||:  || :| ||:|||: :: | ||:| ||| ||||
Db        182 IVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLESLEDTHMIKGEDDALSDKHGCPAY 241

Qy        226 VGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLS 285
              | ||||::  :|||||||||||||| |:|:| | ||| ||:|  || ||||| : :|  :|
Db        242 VSPEILNTTGTYSGKAADVWSLGVMLYTLLVGRYPFHDSDPSALFSKIRRGQFCIPEHVS 301

Qy        286 APARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPD 338
              ||||:| ||||||:|||||  ||||||       |      | :  : |:||:
Db        302 PKARCLIRSLLRREPSERLTAPEILLHPWFEYVLEP-GYVDSEVGTSDQIVPE 353


RESULT 6
Q96RU8
ID   Q96RU8      PRELIMINARY;      PRT;   372 AA.
AC   Q96RU8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   SKIP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kiss-Toth E., Wyllie D.H., Qwarnstrom E.E., Dower S.K.;
RT   "Identification of pro-inflammatory cytokine signalling network
RT   components by transcription expression screening.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF250310; AAK58174.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   372 AA;  40980 MW;  3E2B5C87A4F98FDB CRC64;

  Query Match             39.4%;  Score 744.5;  DB 4;  Length 372;
  Best Local Similarity   46.8%;  Pred. No. 1.8e-54;
  Matches  162;  Conservative   47;  Mismatches  126;  Indels   11;  Gaps    6;

Qy          2 RATPLAAPA-GSLSRKKRLELDD--NLDTERPVQKRARSGP---QPRLPPCLLPLSPPTA 55
              |   |  ||   :  |: |: ||   :  : |      | |    |  ||  |  || |
Db         17 RGPALLFPATRGVPAKRLLDADDAAAVAAKCPRLSECSSPPDYLSPPGSPC-SPQPPPAA 75

Qy         56 PDRATAVATA---SRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYA 112
              |       :|   ||:  |:||   |      |||   || |  |||:|::     : ||
Db         76 PGAGGGSGSAPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKVPPIKHYQDKIRPYI 135

Qy        113 RLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAH 172
              :|| | ::    ||: |    | || :: ||||| |||| |: | ||| ||:|: :|:||
Db        136 QLPSHSNITGIVEVILGETKAYVFFEKSFGDMHSYVRSRKRLREEEAARLFKQIVSAVAH 195

Qy        173 CHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILS 232
              |||  :|| |||| :|||:  || :| ||:|||: :: | ||:| ||| ||||| ||||:
Db        196 CHQSAIVLGDLKLRKFVFSTEERTQLRLESLEDTHIMKGEDDALSDKHGCPAYVSPEILN 255

Qy        233 SRASYSGKAADVWSLGVALPTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLV 292
              :  :|||||||||||||| |:|:| | ||| ||:|  || ||||| : :|  :|  ||||:
Db        256 TTGTYSGKAADVWSLGVMLYTLLVGRYPFHDSDPSALFSKIRRGQFCIPEHISPKARCLI 315

Qy        293 RCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPD 338
              | ||||||:|||||  ||||||      |      | :  :  |:||:
Db        316 RSLLRREPSERLTAPEILLHPWFESVLEP-GYIDSEIGTSDQIVPE 360


RESULT 7
Q9H2Y8
ID   Q9H2Y8      PRELIMINARY;      PRT;   372 AA.
AC   Q9H2Y8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   G-protein-coupled receptor induced protein GIG2.
GN   GIG2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Mayhaus M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,
RA   Nitsch R.M.;
RT   "Identification of a novel nuclear factor Gig2, as an m1-acetylcholine
RT   receptor-induced gene.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF205437; AAG35663.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Receptor; Transferase.
SQ   SEQUENCE   372 AA;  41008 MW;  5F54E50924B1365B CRC64;

  Query Match             39.3%;  Score 742.5;  DB 4;  Length 372;
  Best Local Similarity   46.8%;  Pred. No. 2.6e-54;
  Matches  162;  Conservative   46;  Mismatches  127;  Indels   11;  Gaps    6;

Qy          2 RATPLAAPA-GSLSRKKRLELDD--NLDTERPVQKRARSGP---QPRLPPCLLPLSPPTA 55
              |   |  ||   :  |: |: ||   :  : |      | |    |  ||  |  || |
Db         17 RGPALLFPATRGVPAKRLLDADDAAAVAAKCPRLSECSSPPDYLSPPGSPC-SPQPPPAA 75

Qy         56 PDRATAVATA---SRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYA 112
              |       :|   ||:  |:||   |      |||   || |  |||:|::     : ||
Db         76 PGAGGGSGSAPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKVPPIKHYQDKIRPYI 135

Qy        113 RLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAH 172
              :|| | ::    ||: |    | || :: ||||| |||| |: | ||| ||:|: :|:||
Db        136 QLPSHSNITGIVEVILGETKAYVFFEKDFGDMHSYVRSRKRLREEEAARLFKQIVSAVAH 195

Qy        173 CHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILS 232
              |||  :|| |||| :|||:  || :| ||:|||: :: | ||:| ||| ||||| ||||:
Db        196 CHQSAIVLGDLKLRKFVFSTEERTQLRLESLEDTHIMKGEDDALSDKHGCPAYVSPEILN 255

Qy        233 SRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLV 292
              :  :|||||||||||||| |:|:| | ||| ||:|  || ||||| : :|  :|  ||||:
Db        256 TTGTYSGKAADVWSLGVMLYTLLVGRYPFHDSDPSALFSKIRRGQFCIPEHISPKARCLI 315

Qy        293 RCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPD 338
              | ||||||:|||||  ||||||      |      | :  :  |:||:
Db        316 RSLLRREPSERLTAPEILLHPWFESVLEP-GYIDSEIGTSDQIVPE 360


RESULT 8
Q91W04
ID   Q91W04      PRELIMINARY;      PRT;   372 AA.
AC   Q91W04;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to phosphoprotein regulated by mitogenic pathways.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Skin;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; BC006800; AAH06800.1; -.
DR   EMBL; AK028626; BAC26038.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   372 AA;  41281 MW;  AD29BB4E640B4B62 CRC64;

  Query Match             39.3%;  Score 742.5;  DB 11;  Length 372;
  Best Local Similarity   46.7%;  Pred. No. 2.6e-54;
  Matches  165;  Conservative   44;  Mismatches  119;  Indels   25;  Gaps    6;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLP---LSPPTAP-- 56
              |   |  ||   :  |||     |||:     | :  |||  |  |   |||| :|
Db         17 RGPGLLFPAARGTPAKRL-----LDTD---DAGAVAAKCPRLSECSSPPDYLSPPGSPCS 68

Qy         57 ------DRATAVATASRLGP-----YVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
```

```
                  : |  :  |  ||     |:||   |     |||   || |  || :|::
Db         69 PQPPPSTQGTGGSCVSSPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKEFPIKHYQ 128

Qy        106 AVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQ 165
              : || :|| | ::    ||| |    | || :  ||||| |||| |: | ||| ||:|
Db        129 DKIRPYIQLPSHSNITGIVEVLLGESKAYVPFEKDFGDMHSYVRSRKRLREEEAARLFKQ 188

Qy        166 MATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAY 225
              : :|:|||||  :|| |||| :|||:  || :| ||:|||: :: | ||:| ||| ||||
Db        189 IVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLESLEDTHIIKGEDDALSDKHGCPAY 248

Qy        226 VGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLS 285
              | ||||::  :||||||||||||| |:|:| | ||| ||:|  || ||||| : :|  :|
Db        249 VSPEILNTTGTYSGKAADVWSLGVMLYTLLVGRYPFHDSDPSALFSKIRRGQFCIPEHVS 308

Qy        286 APARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPD 338
              ||||:| ||||||:|||||  ||||||      |     | :  : |:||:
Db        309 PKARCLIRSLLRREPSERLTAPQILLHPWFEYVLEP-GYVDSEIGTSDQIVPE 360


RESULT 9
Q8K4K4
ID   Q8K4K4      PRELIMINARY;      PRT;   372 AA.
AC   Q8K4K4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   TRB-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kiss-Toth E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA   Bagstaff S.M., Wyllie D.H., Harte M., O'Neill L.A.J., Qwarnstrom E.E.,
RA   Dower S.K.;
RT   "Mammalian homologs of Drosophila tribbles (htrb) control mitogen
RT   activated protein kinase signaling.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF358866; AAM45478.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   372 AA;  41282 MW;  3A3DE82B46CD907F CRC64;

  Query Match             38.6%;  Score 729.5;  DB 11;  Length 372;
  Best Local Similarity   46.2%;  Pred. No. 3.3e-53;
  Matches  163;  Conservative   44;  Mismatches  121;  Indels   25;  Gaps    6;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLP---LSPPTAP-- 56
              |   |  ||   :  |||     |||:      | :   |||  |  |   |||| :|
Db         17 RGPGLLFPAARGTPAKRL-----LDTD---DAGAVAAKCPRLSECSSPPDYLSPPGSPCS 68

Qy         57 ------DRATAVATASRLGP-----YVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
                  :  |  :  |  ||      |:||   |      |||   || |  || :|::
Db         69 PQPPPSTQGTGGSCVSSPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKEFPIKHYQ 128

Qy        106 AVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQ 165
              : || :|| | ::    ||| |    | || :  ||||| |||| |: | ||| ||:|
Db        129 DKIRPYIQLPSHSNITGIVEVLLGESKAYVPFEKDFGDMHSYVRSRKRLREEEAARLFKQ 188

Qy        166 MATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAY 225
              : :|:|||||  :|| |||| :|||:  || :| | :|||: :: | ||:| ||| ||||
Db        189 IVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLGSLEDTHIIKGEDDALSDKHGCPAY 248

Qy        226 VGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLS 285
              | ||||::  :|||||||||||| |:|:  | ||| ||:|  || ||||| : :|  :|
Db        249 VSPEILNTTGTYSGKAADVWSLGVMLYTLWVGRYPFHDSDPSALFSKIRRGQFCIPEHVS 308

Qy        286 APARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPD 338
              ||||:| ||||||:|||||  ||||||      |     | :  : |:||:
Db        309 PKARCLIRSLLRREPSERLTAPQILLHPWFEYVLEP-GYVDSEIGTSDQIVPE 360


RESULT 10
Q8R2V8
ID   Q8R2V8      PRELIMINARY;      PRT;   257 AA.
AC   Q8R2V8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 29.2 kDa protein (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC027159; AAH27159.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hypothetical protein; ATP-binding; Transferase.
FT   NON_TER       1       1
SQ   SEQUENCE   257 AA;  29232 MW;  84EE4C0DD4767F51 CRC64;

  Query Match             34.5%;  Score 653;  DB 11;  Length 257;
  Best Local Similarity   52.3%;  Pred. No. 5.7e-47;
  Matches  134;  Conservative   34;  Mismatches   68;  Indels   20;  Gaps    3;

Qy         96 CKVYPV---QEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRH 152
              |||: :   ||:||   |   |  | :: : ||:| |     | || |::||||| ||:
Db          3 CKVFEISCYQESLA---PCFCLSAHSNINQITEILLGETKAYVPFERSYGDMHSFVRTCK 59

Qy        153 RIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGP 212
              :: | ||| || |:|:|:|||| |||||||| :|:| | || :: ||:|||: :| |
Db         60 KLREEEAARLFYQIASAVAHCHDGGLVLRDLKLRKFIFKDEERTRVKLESLEDAYILRGD 119

Qy        213 DDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGK 272
              |||| ||| ||||| ||||::  |||||||||||||| |:||| | ||| | ||  || |
Db        120 DDSLSDKHGCPAYVSPEILNTSGSYSGKAADVWSLGVMLYTMLVGRYPFHDIEPSSLFSK 179

Qy        273 IRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEA 332
              |||| : :|  ||  |:||:| :|||||:||||:  || |||   |
Db        180 IRRGQFNIPETLSPKAKCLIRSILRREPSERLTSQEILDHPWFSTD------------- 225

Qy        333 AQVVPDGLGLDEAREE 348
                |  | |  || ::
Db        226 FSVSNSGFGAKEACDQ 241


RESULT 11
O15180
ID   O15180      PRELIMINARY;      PRT;   224 AA.
AC   O15180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Phosphoprotein (Fragment).
GN   C8FW.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RA   Wilkain F.;
```

```
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RX   MEDLINE=98000262; PubMed=9342215;
RA   Wilkin F., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,
RA   Dumont J.E., Maenhaut C.;
RT   "Characterization of a phosphoprotein whose mRNA is regulated by the
RT   mitogenic pathways in dog thyroid cells.";
RL   Eur. J. Biochem. 248:660-669(1997).
DR   EMBL; AJ000480; CAA04119.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
FT   NON_TER      1      1
SQ   SEQUENCE   224 AA;  25479 MW;  740CD0905F86499B CRC64;

  Query Match             32.3%;  Score 611.5;  DB 4;  Length 224;
  Best Local Similarity   56.7%;  Pred. No. 1.5e-43;
  Matches  119;  Conservative   31;  Mismatches   59;  Indels    1;  Gaps    1;

Qy        129 GTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRF 188
              |     | || :: ||||| |||| |: | ||| ||:|: :|:|||||  :|| |||| :|
Db          4 GETKAYVFFEKSFGDMHSYVRSRKRLREEEAARLFKQIVSAVAHCHQSAIVLGDLKLRKF 63

Qy        189 VFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLG 248
              ||:  || :| ||:|||: :: | ||:| ||| ||||| ||||::  :|||||||||||
Db         64 VFSTEERTQLRLESLEDTHIMKGEDDALSDKHGCPAYVSPEILNTTGTYSGKAADVWSLG 123

Qy        249 VALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATG 308
              | |:|:| | ||| ||:|  || ||||| : :|  :|  ||||:| ||||||:|||||
Db        124 VMLYTLLVGRYPFHDSDPSALFSKIRRGQFCIPEHISPKARCLIRSLLRREPSERLTAPE 183

Qy        309 ILLHPWLRQDPMPLAPTRSHLWEAAQVVPD 338
              ||||||      |     | :  : |:||:
Db        184 ILLHPWFESVLEP-GYIDSEIGTSDQIVPE 212


RESULT 12
Q9V3Z1
ID   Q9V3Z1      PRELIMINARY;      PRT;   484 AA.
AC   Q9V3Z1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative TRBL protein (CG5408 protein).
GN   TRBL OR CG5408.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Seher T.C., Leptin M.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003591; AAF51590.1; -.
DR   EMBL; AF204688; AAF26374.1; -.
DR   FlyBase; FBgn0028978; trbl.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hypothetical protein; ATP-binding; Transferase.
SQ   SEQUENCE   484 AA;  54077 MW;  3E3B1D3B5645B0D7 CRC64;

  Query Match             23.0%;  Score 435;  DB 5;  Length 484;
  Best Local Similarity   33.4%;  Pred. No. 2.5e-28;
  Matches  116;  Conservative   45;  Mismatches  134;  Indels   52;  Gaps    8;

Qy         31 VQKRARSGPQPRLPPCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYRALHCPT 90
              :|:|     ||       :    | | |     |||  |                | :   |
Db        104 IQQRYLISAQPSHISAAVAAKTP-ASYRHLVDLTASNL--------------RCVDIFT 147

Qy         91 GTEYTCKVYPVQEAL-AVLEPYARLPPHKHVARPTEVLA---------------GTQLL 133
              | :: |::  | | |  |   | :|  |    | : :                 | :|
Db        148 GEQFLCRI--VNEPLHKVQRAYFQLQQHDEELRRSTIYGHPLIRPVHDIIPLTKDRTYIL 205

Qy        134 YA----------PFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDL 183
              |           |  : ::|: :|   |: | ||  :| |:   :  ||::|::||||
Db        206 IAPVPQERDSTGGVTGVYENLHTYIRHAKRLCETEARAIFHQICQTVQVCHRNGIILRDL 265

Qy        184 KLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAAD 243
              || || | |  | ||  |:|| | :| | ||:| ||  || |  ||:|  : :| || ||
Db        266 KLKRFYPIDEARTKLQYESLEGSMILDGEDDTLSDKIGCPLYTAPELLCPQQTYKGKPAD 325

Qy        244 VWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAER 303
              :|||||  |:||| | ||| :      |   || |   :|  ||   | |:  |||::  ||
Db        326 MWSLGVILYTMLVGQYPFYEKANCNLITVIRHGNVQIPLTLSKSVRWLLLSLLRKDYTER 385

Qy        304 LTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARBEEG 350
              :||: | | ||||:      |   :|    :|  |     :| |:||
Db        386 MTASHIFLTPWLREQ----RPFHMYLPVDVEVAED---WSDAEEDEG 425


RESULT 13
Q8BFS7
ID   Q8BFS7      PRELIMINARY;      PRT;   218 AA.
AC   Q8BFS7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE   Similar to phosphoprotein regulated BY mitogenic PATHWAYS
DE   homolog.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK040738; BAC30688.1; -.
DR   EMBL; AK041212; BAC30865.1; -.
SQ   SEQUENCE   218 AA;  23907 MW;  F3A5AFE16E3450A8 CRC64;

  Query Match             17.2%;  Score 326;  DB 11;  Length 218;
  Best Local Similarity   40.0%;  Pred. No. 1.4e-19;
  Matches   84;  Conservative   23;  Mismatches   79;  Indels   24;  Gaps    5;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLP---LSPPTAP-- 56
              |   |  ||   :  |||     |||:      | :   |||  |  |   |||| :|
Db         17 RGPGLLFPAARGTPAKRL-----LDTD---DAGAVAAKCPRLSECSSPPDYLSPPGSPCS 68

Qy         57 ------DRATAVATASRLGP-----YVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
                    : |  :  |  ||     |:||   |      |||   || |  || :|::
Db         69 PQPPPSTQGTGGSCVSSPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKEPPIKHYQ 128

Qy        106 AVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQ 165
              : || :|| | ::    ||| |    | || :  ||||| |||| |: | ||| ||:|
Db        129 DKIRPYIQLPSHSNITGIVEVLLGESKAYVFFEKDFGDMHSYVRSRKRLREEEAARLFKQ 188

Qy        166 MATALAHCHQHGLVLRDLKLCRFVFADRER 195
              : :|:|||||  :|| |||| :|||:  ||
Db        189 IVSAVAHCHQSAIVLGDLKLRKFVFSTEER 218


RESULT 14
Q8BJR9
ID   Q8BJR9      PRELIMINARY;      PRT;   379 AA.
AC   Q8BJR9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to phosphoprotein regulated BY mitogenic PATHWAYS homolog
DE   (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK080228; BAC37854.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   379 AA;  40469 MW;  4A656194F5EB6BE4 CRC64;

  Query Match             17.2%;  Score 326;  DB 11;  Length 379;
  Best Local Similarity   40.0%;  Pred. No. 2.8e-19;
  Matches   84;  Conservative   23;  Mismatches   79;  Indels   24;  Gaps    5;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLP---LSPPTAP-- 56
              |   |  ||   :  |||     |||:      | :   |||  |  |   |||| :|
Db        178 RGPGLLFPAARGTPAKRL-----LDTD---DAGAVAAKCPRLSECSSPPDYLSPPGSPCS 229

Qy         57 ------DRATAVATASRLGP-----YVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
                    : |  :  |  ||     |:||   |      |||   || |  || :|::
Db        230 PQPPPSTQGTGGSCVSSPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKEPPIKHYQ 289

Qy        106 AVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQ 165
              : || :|| | ::    ||| |    | || :  ||||| |||| |: | ||| ||:|
Db        290 DKIRPYIQLPSHSNITGIVEVLLGESKAYVFFEKDFGDMHSYVRSRKRLREEEAARLFKQ 349

Qy        166 MATALAHCHQHGLVLRDLKLCRFVFADRER 195
              : :|:|||||  :|| |||| :|||:  ||
Db        350 IVSAVAHCHQSAIVLGDLKLRKFVFSTEER 379


RESULT 15
Q8BZX3
ID   Q8BZX3      PRELIMINARY;      PRT;   388 AA.
AC   Q8BZX3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to phosphoprotein regulated BY mitogenic PATHWAYS homolog
DE   (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK033358; BAC28245.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   388 AA;  41477 MW;  6DC664A3468E2BCB CRC64;

  Query Match             17.2%;  Score 326;  DB 11;  Length 388;
  Best Local Similarity   40.0%;  Pred. No. 2.8e-19;
  Matches   84;  Conservative   23;  Mismatches   79;  Indels   24;  Gaps    5;

Qy          2 RATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLP---LSPPTAP-- 56
              |   |  ||   :  |||     |||:      | :   |||  |  |   |||| :|
Db        187 RGPGLLFPAARGTPAKRL-----LDTD---DAGAVAAKCPRLSECSSPPDYLSPPGSPCS 238

Qy         57 ------DRATAVATASRLGP-----YVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
                    : |  :  |  ||     |:||   |      |||   || |  || :|::
Db        239 PQPPPSTQGTGGSCVSSPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKEPPIKHYQ 298

Qy        106 AVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQ 165
              : || :|| | ::    ||| |    | || :  ||||| |||| |: | ||| ||:|
Db        299 DKIRPYIQLPSHSNITGIVEVLLGESKAYVFFEKDFGDMHSYVRSRKRLREEEAARLFKQ 358

Qy        166 MATALAHCHQHGLVLRDLKLCRFVFADRER 195
              : :|:|||||  :|| |||| :|||:  ||
Db        359 IVSAVAHCHQSAIVLGDLKLRKFVFSTEER 388


Search completed: January 15, 2004, 14:58:44
Job time : 43 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:55:46 ; Search time 18 Seconds
(without alignments)
935.309 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 1891
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1887	99.8	358	1	NIPK_HUMAN	Q96ru7 homo sapien
2	1345	71.1	354	1	NIPK_MOUSE	Q8k4k2 mus musculu
3	1331.5	70.4	349	1	NIPK_RAT	Q9wtq6 rattus norv
4	309	16.3	619	1	SNF1_CANTR	O94168 candida tro
5	304.5	16.1	646	1	CNK_HUMAN	Q9h4b4 homo sapien
6	304	16.1	620	1	SNF1_CANAL	P52497 candida alb
7	298.5	15.8	631	1	CNK_MOUSE	Q60806 mus musculu
8	291	15.4	615	1	CNK_RAT	Q9r011 rattus norv
9	287	15.2	1518	1	KKK1_YEAST	P34244 saccharomyc
10	284.5	15.0	611	1	SNF1_CANGA	Q00372 candida gla
11	281	14.9	512	1	KI10_ARATH	Q38997 arabidopsis
12	274.5	14.5	633	1	SNF1_YEAST	P06782 saccharomyc
13	273.5	14.5	533	1	KCCD_RAT	P15791 rattus norv
14	271.5	14.4	593	1	CDR1_SCHPO	P07334 schizosacch
15	271	14.3	576	1	SNF1_SCHPO	O74536 schizosacch
16	269	14.2	542	1	KCCB_MOUSE	P28652 mus musculu
17	268	14.2	622	1	YNA3_CAEEL	P45894 caenorhabdi
18	266.5	14.1	499	1	KCCD_HUMAN	Q13557 homo sapien
19	266	14.1	542	1	KCCB_RAT	P08413 rattus norv
20	266	14.1	664	1	KCCB_HUMAN	Q13554 homo sapien
21	263	13.9	1142	1	GIN4_YEAST	Q12263 saccharomyc
22	261.5	13.8	552	1	AAK2_HUMAN	P54646 homo sapien
23	261.5	13.8	891	1	KIN1_SCHPO	P22987 schizosacch
24	260.5	13.8	552	1	AAK2_RAT	Q09137 rattus norv
25	260.5	13.8	786	1	SN1L_HUMAN	P57059 homo sapien
26	258	13.6	502	1	RKI1_SECCE	Q02723 secale cere
27	256	13.5	610	1	CDP1_ARATH	Q06850 arabidopsis
28	255.5	13.5	478	1	KCCA_HUMAN	Q9uqm7 homo sapien
29	255.5	13.5	478	1	KCCA_RAT	P11275 rattus norv
30	252.5	13.4	433	1	STKB_HUMAN	Q15831 homo sapien
31	250	13.2	735	1	K6A1_RAT	Q63531 rattus norv
32	247.5	13.1	685	1	SNK_HUMAN	Q9nyy3 homo sapien
33	247	13.1	332	1	AAIP_WHEAT	Q02066 triticum ae
34	246.5	13.0	724	1	K6A1_MOUSE	P18653 mus musculu
35	244.5	12.9	478	1	KCCA_MOUSE	P11798 mus musculu
36	244.5	12.9	735	1	K6A1_HUMAN	Q15418 homo sapien
37	243.5	12.9	472	1	KCCG_HUMAN	Q13555 homo sapien
38	243.5	12.9	527	1	KCCG_RAT	P11730 rattus norv
39	243.5	12.9	529	1	KCCG_MOUSE	Q923t9 mus musculu
40	242.5	12.8	353	1	ASK2_ARATH	P43292 arabidopsis
41	242	12.8	661	1	Y537_HUMAN	O60285 homo sapien
42	242	12.8	915	1	KCC4_YEAST	P25389 saccharomyc
43	241	12.7	542	1	CDP3_ORYSA	P53684 oryza sativ
44	240.5	12.7	548	1	AAK1_RAT	P54645 rattus norv
45	240	12.7	682	1	SNK_MOUSE	P53351 mus musculu

ALIGNMENTS

RESULT 1
NIPK_HUMAN
ID NIPK_HUMAN STANDARD; PRT; 358 AA.
AC Q96RU7; Q8TAI6; Q9H5M8; Q9NUD2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neuronal cell death inducible putative kinase (SKIP3).
GN NIPK OR C20ORF97.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiss-Toth E., Wyllie D.H., Qwarnstrom E.E., Dower S.K.;
RT "Identification of pro-inflammatory cytokine signalling network
RT components by transcription expression screening.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).

```
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cervix, and Muscle;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May play an important role in a common pathway leading
CC       to programmed neuronal cell death. Does not appear to function in
CC       the programmed death of non-neuronal cells. May serve as an
CC       endogenous antagonist competing for substrate with functional
CC       kinases that act to promote neuronal cell survival (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 protein kinase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF250311; AAK58175.1; -.
DR   EMBL; AK026945; BAB15597.1; -.
DR   EMBL; AL034548; CAB81634.1; -.
DR   EMBL; BC019363; AAH19363.1; -.
DR   EMBL; BC027484; AAH27484.1; -.
DR   Genew; HGNC:16228; C20orf97.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Apoptosis.
FT   DOMAIN       68    316       PROTEIN KINASE.
FT   CONFLICT     84     84       Q -> R (IN REF. 4; AAH27484).
FT   CONFLICT    105    105       L -> P (IN REF. 2).
FT   CONFLICT    114    114       L -> V (IN REF. 1).
FT   CONFLICT    194    195       ER -> DREK (IN REF. 1).
SQ   SEQUENCE   358 AA;  39577 MW;  CE15FD89A81E8D63 CRC64;

  Query Match           99.8%;  Score 1887;  DB 1;  Length 358;
  Best Local Similarity   99.7%;  Pred. No. 1e-140;
  Matches  357;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         61 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300

Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358


RESULT 2
NIPK_MOUSE
ID   NIPK_MOUSE     STANDARD;      PRT;   354 AA.
AC   Q8K4K2; Q921E7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Neuronal cell death inducible putative kinase (TRB-3).
GN   NIPK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kiss-Toth E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA   Bagstaff S.M., Wyllie D.H., Harte M., O'Neill L.A.J., Qwarnstrom E.E.,
RA   Dower S.K.;
RT   "Mammalian homologs of Drosophila tribbles (htrb) control mitogen
RT   activated protein kinase signaling.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
```

```
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May play an important role in a common pathway leading
CC       to programmed neuronal cell death. Does not appear to function in
CC       the programmed death of non-neuronal cells. May serve as an
CC       endogenous antagonist competing for substrate with functional
CC       kinases that act to promote neuronal cell survival (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 protein kinase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF358868; AAM45476.1; -.
DR   EMBL; AK089931; BAC41002.1; -.
DR   EMBL; BC012955; AAH12955.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Apoptosis.
FT   DOMAIN       68    315       PROTEIN KINASE.
FT   CONFLICT    157    157       S -> P (IN REF. 3).
FT   CONFLICT    219    219       K -> T (IN REF. 1).
FT   CONFLICT    239    264       MISSING (IN REF. 3).
FT   CONFLICT    301    354       SERLVALGILLHPWLREDHGRVSPPQSDRREMDQVVPDGPQ
FT                                LEEAEEGEVGLYG -> CRATCGPGNPLASLVERGSRPSLS
FT                                STV (IN REF. 2).
SQ   SEQUENCE   354 AA;  39022 MW;  2CB283FC119F859F CRC64;

  Query Match             71.1%;  Score 1345;  DB 1;  Length 354;
  Best Local Similarity   74.0%;  Pred. No. 3e-98;
  Matches  265;  Conservative   29;  Mismatches   60;  Indels    4;  Gaps    2;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              |||||||| |    ||| || |||:| : || || |  |:|   | ||| ||| | | :
Db          1 MRATPLAASADVSCRKKPLEFDDNIDAKCPVLKRVRDEPEPGPLPSLLPPSPPPASDLSP 60

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||| |:|||||:||| |:|  :||||||||||||||||||  || ||| |||||| |:||
Db         61 AVAPATRLGPYILLEREQGSCSYRALHCPTGTEYTCKVYPASEAQAVLAPYARLPTHQHV 120

Qy        121 ARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||| |::||| |||:||||:|||||||  ||| ||| ||||||:|:||||:|||||
Db        121 ARPTEVLLGSRLLYIPFTKTHGDLHSLVRSRRGIPESEAAGLFRQMASAVAHCHKHGLVL 180

Qy        181 RDLKLCRPVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||| ||||:: || |||||||||:||:|| ||||||||||||||||||||| |||||
Db        181 RDLKLRRFVFSNCERTKLVLENLEDACVMTGSDDSLWDKHACPAYVGPEILSSRPSYSGK 240

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              |||||||||||||||| ||| |||||||||||||| :||| |||||||||:|||||:||
Db        241 AADVWSLGVALFTMLAGRYPFHDSEPVLLFGKIRRGTFALPEGLSAPARCLIRCLLRKEP 300

Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              :||| | |||||||||:|   ::| :|   |  ||||||  |:||  |||  || |||
Db        301 SERLVALGILLHPWLREDHGRVSPPQSDRREMDQVVPDGPQLEEA--EEG--EVGLYG 354


RESULT 3
NIPK_RAT
ID   NIPK_RAT       STANDARD;      PRT;   349 AA.
AC   Q9WTQ6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neuronal cell death inducible putative kinase.
GN   NIPK.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Neuron;
RX   MEDLINE=99262087; PubMed=10329375;
RA   Mayumi-Matsuda K., Kojima S., Suzuki H., Sakata T.;
RT   "Identification of a novel kinase-like gene induced during neuronal
RT   cell death.";
RL   Biochem. Biophys. Res. Commun. 258:260-264(1999).
CC   -!- FUNCTION: May play an important role in a common pathway leading
CC       to programmed neuronal cell death. Does not appear to function
CC       in the programmed death of non-neuronal cells. May serve as an
CC       endogenous antagonist competing for substrate with functional
CC       kinases that act to promote neuronal cell survival.
CC   -!- TISSUE SPECIFICITY: Detected only in the lung. Not detected in the
CC       heart, brain, spleen, liver, skeletal muscle, kidney and testis.
CC   -!- INDUCTION: Expression induced during programmed cell death evoked
CC       in neuronal cells by NGF-depletion.
CC   -!- SIMILARITY: Contains 1 protein kinase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB020967; BAA77582.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Apoptosis.
FT   DOMAIN       63    310       PROTEIN KINASE.
SQ   SEQUENCE   349 AA;  38602 MW;  3050F9BFB946D815 CRC64;

  Query Match             70.4%;  Score 1331.5;  DB 1;  Length 349;
  Best Local Similarity   73.5%;  Pred. No. 3.4e-97;
  Matches  263;  Conservative   26;  Mismatches   60;  Indels    9;  Gaps    3;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              |||| ||| |    ||| || |||:| | || || |  |:|   | |     | | | :
Db          1 MRATSLAASADVPCRKKPLEFDDNIDVECPVLKRVRDEPEPGPTPSL-----PPASDLSP 55

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||| |:|||||:||| |:|   |||||||||||||||||  || ||| ||||| |:||
Db         56 AVAPATRLGPYILLEREQGNCTYRALHCPTGTEYTCKVYPASEAQAVLAPYARLPTHQHV 115

Qy        121 ARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||| |:|||| |||:||||:|||||||  ||||||| ||||||:|:||||:|||:|
Db        116 ARPTEVLLGSQLLYTPFTKTHGDLHSLVRSRRGIPEPEAAALFRQMASAVAHCHKHGLIL 175
```

```
Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||| ||||:: || ||||||||||:||:||||||||||||||||||||||| ||||:
Db        176 RDLKLRRFVFSNCERTKLVLENLEDACVMTGPDDSLWDKHACPAYVGPEILSSRPSYSGR 235

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||||||||||||||||| |||||||| |||||||| :||| |||| |||||||||||||
Db        236 AADVWSLGVALFTMLAGRYPFQDSEPALLFGKIRRGTFALPEGLSASARCLIRCLLRREP 295

Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              :||| | |||||||||:|   ::| ||    |  |||||  |:||  |||  || |||
Db        296 SERLVALGILLHPWLREDCSQVSPPRSDRREMDQVVPDGPQLEEA--EEG--EVGLYG 349


RESULT 4
SNF1_CANTR
ID   SNF1_CANTR     STANDARD;      PRT;   619 AA.
AC   O94168;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN   SNF1.
OS   Candida tropicalis (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5482;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kanai T., Ogawa K., Ueda M., Tanaka A.;
RT   "Genetic evaluation of the function of SNF1 in Candida tropicalis.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
CC       INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
CC       PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       SNF1 SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB024535; BAA75889.1; -.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Phosphorylation; Carbohydrate metabolism; Nuclear protein.
FT   DOMAIN       17     29       POLY-HIS.
FT   DOMAIN       52    303       PROTEIN KINASE.
FT   NP_BIND      58     66       ATP (BY SIMILARITY).
FT   BINDING      81     81       ATP (BY SIMILARITY).
FT   ACT_SITE    174    174       BY SIMILARITY.
FT   MOD_RES     207    207       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ   SEQUENCE   619 AA;  70323 MW;  0FCF1FC3DCE706D7 CRC64;

  Query Match             16.3%;  Score 309;  DB 1;  Length 619;
  Best Local Similarity   30.9%;  Pred. No. 8.5e-17;
  Matches   95;  Conservative   48;  Mismatches  128;  Indels   36;  Gaps   12;

Qy         40 QPRLPPCLLPLSPPTAPDRATAVATASRLGPYVLLEP-EEG--GRAYRALHCPTGTEYTC 96
              ||  |   :|: |   |        |:|:| | :::   ||  |:   | |  || :
Db         32 QPAQP---IPIDPNVNP--------ANRIGRYQIIKTLGEGSFGKVKLAQHVGTGQKVAL 80

Qy         97 KVYPVQEALAVLE---------PYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDMHSL 147
              |:   :: ||  :          | ||  | |: :  :|:     :           ::
Db         81 KIIN-RKTLAKSDMQGRVEREISYLRLLRHPHIIKLYDVIKSKDEIIMVIEFAGKELFDY 139

Qy        148 VRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSC 207
              :  | ::|| ||   |:|:  |: :||:| :| ||||    :  |:   |:    |  |
Db        140 IVQRGKMPEDEARRFFQQIIAAVEYCHRHKIVHRDLKPENLLLDDQLNVKIADFGL--SN 197

Qy        208 VLTGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSE 265
              ::|   |  : | :|  | |  ||::| :  |:|   |||| || |: || |  || |
Db        198 IMT---DGNFLKTSCGSPNYAAPEVISGKL-YAGPEVDVWSSGVILYVMLCGRLPFDDEF 253

Qy        266 PVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMP---L 322
              || ||  | | ||  ||  |: |:  :|   |  |:|   |:   | :|| ||   |
Db        254 IPALFKKISNGVYTLPNYLSPGAKHLLTRMLVVNPLNRITIHEIMEDEWFKQD-MPDYLL 312

Qy        323 APTRSHL 329
                |  | :
Db        313 PPDLSKI 319


RESULT 5
CNK_HUMAN
ID   CNK_HUMAN      STANDARD;      PRT;   646 AA.
AC   Q9H4B4; Q15767;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-
DE   inducible kinase) (Proliferation-related kinase).
GN   CNK OR FNK OR PRK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=20493044; PubMed=11039900;
RA   Holtrich U., Wolf G., Yuan J., Bereiter-Hahn J., Karn T., Weiler M.,
RA   Kauselmann G., Rehli M., Andreesen R., Kaufmann M., Kuhl D.,
RA   Strebhardt K.;
RT   "Adhesion induced expression of the serine/threonine kinase Fnk in
RT   human macrophages.";
RL   Oncogene 19:4832-4839(2000).
RN   [2]
RP   SEQUENCE OF 28-646 FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96325053; PubMed=8702627;
RA   Li B., Ouyang B., Pan H., Reissmann P.T., Slamon D.J., Arceci R.,
RA   Lu L., Dai W.;
RT   "Prk, a cytokine-inducible human protein serine/threonine kinase whose
RT   expression appears to be down-regulated in lung carcinomas.";
RL   J. Biol. Chem. 271:19402-19408(1996).
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=98019242; PubMed=9353331;
RA   Ouyang B., Pan H., Lu L., Li J., Stambrook P., Li B., Dai W.;
RT   "Human Prk is a conserved protein serine/threonine kinase involved in
RT   regulating M phase functions.";
RL   J. Biol. Chem. 272:28646-28651(1997).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=20027391; PubMed=10557092;
RA   Ouyang B., Li W., Pan H., Meadows J., Hoffmann I., Dai W.;
RT   "The physical association and phosphorylation of Cdc25C protein
RT   phosphatase by Prk.";
RL   Oncogene 18:6029-6036(1999).
CC   -!- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN REGULATING M
CC       PHASE FUNCTIONS DURING THE CELL CYCLE. MAY ALSO BE PART OF THE
```

```
CC       SIGNALING NETWORK CONTROLLING CELLULAR ADHESION. IN VITRO, IS ABLE
CC       TO PHOSPHORYLATE CDC25C AND CASEIN.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: BINDS TO THE CALCIUM/INTEGRIN-BINDING PROTEIN (CIB). THIS
CC       INTERACTION PROBABLY OCCURS VIA THE POLO-BOX DOMAIN.
CC   -!- SUBCELLULAR LOCATION: Membrane-associated.
CC   -!- TISSUE SPECIFICITY: TRANSCRIPTS ARE HIGHLY DETECTED IN PLACENTA,
CC       LUNG, FOLLOWED BY SKELETAL MUSCLE, HEART, PANCREAS, OVARIES AND
CC       KIDNEY AND WEAKLY DETECTED IN LIVER AND BRAIN. MAY HAVE A SHORT
CC       HALF-LIVE. IN CELLS OF HEMATOPOIETIC ORIGIN, STRONGLY AND
CC       EXCLUSIVELY DETECTED IN TERMINALLY DIFFERENTIATED MACROPHAGES.
CC       TRANSCRIPTS EXPRESSION APPEARS TO BE DOWN-REGULATED IN PRIMARY
CC       LUNG TUMOR.
CC   -!- INDUCTION: CYTOKINE AND CELLULAR ADHESION TRIGGER FNK INDUCTION.
CC   -!- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
CC       CELLS EXIT MITOSIS (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       CDC5/POLO SUBFAMILY.
CC   -!- SIMILARITY: Contains 2 POLO box domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ293866; CAC10659.1; -.
DR   EMBL; U56998; AAC50637.1; ALT_INIT.
DR   Genew; HGNC:2154; CNK.
DR   GK; Q9H4B4; -.
DR   MIM; 602913; -.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR   GO; GO:0000074; P:regulation of cell cycle; TAS.
DR   InterPro; IPR000959; POLO_box.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00659; POLO_box; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50078; POLO_BOX; 2.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW   Phosphorylation.
FT   DOMAIN       62    314       PROTEIN KINASE.
FT   NP_BIND      68     76       ATP (BY SIMILARITY).
FT   BINDING      91     91       ATP (BY SIMILARITY).
FT   ACT_SITE    185    185       BY SIMILARITY.
FT   DOMAIN      470    537       POLO BOX 1.
FT   DOMAIN      567    637       POLO BOX 2.
FT   CONFLICT     99     99       V -> A (IN REF. 2).
FT   CONFLICT    353    353       V -> G (IN REF. 2).
FT   CONFLICT    419    419       H -> D (IN REF. 2).
FT   CONFLICT    464    470       PSEWVGF -> VSKWVDY (IN REF. 2).
FT   CONFLICT    522    522       R -> P (IN REF. 2).
SQ   SEQUENCE   646 AA;  71789 MW;  C20147CD0FC8A3B4 CRC64;

  Query Match             16.1%;  Score 304.5;  DB 1;  Length 646;
  Best Local Similarity   28.9%;  Pred. No. 2e-16;
  Matches  101;  Conservative   48;  Mismatches  134;  Indels   67;  Gaps   15;

Qy         29 RPVQKRAR-----SGPQPRLPPCL-----------LPLSPP----TAPDRATAVATASRL 68
              || |: |      :|| |  ||              || | |    | |           |
Db         11 RPFQRTAAATAPPAGPGP--PPSALRGPELEMLAGLPTSDPGRLITDPRSGRTYLKGRLL 68

Qy         69 GPYVLLEPEEGG--RAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLP---------- 115
              |       :||  | | |    ||: |  || |      |::|: |
Db         69 G--------KGGPARCYEATDTETGSAYAVKVIPQSR---VVKPHQREKILNEIELHRDL 117

Qy        116 PHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCH 174
              |:|: | :          :| |    :   :  : ::|| : |||     ||: : | : |
Db        118 QHRHIVRFSHHFEDADNIYIFLELCSRKSLAHIWKARHTLLEPEVRYYLRQILSGLKYLH 177

Qy        175 QHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHAC--PAYVGPEILS 232
              | |:: ||||  |      | :| : :   :  | |:     |  |  | || ||:|
Db        178 QRGILHRDLKLGNFFIT--ENMELKVGDFGLAARLEPPEQR--KKTICGTPNYVAPEVLL 233

Qy        233 SRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLV 292
              :    |  |||||| ::|:| |  ||: ::    :  |::  | ||| || ||| |:
Db        234 RQG--HGPEADVWSLGCVMYTLLCGSPPFETADLKETYRCIKQVHYTLPASLSLPARQLL 291

Qy        293 RCLLRREPAERLTATGILLHPWLRQ----DPMPLAPTRSHLWEAAQVVPD 338
                :||  | :| :   || | :  :    | :|::        :   |||
Db        292 AAILRASPRDRPSIDQILRHDFFTKGYTPDRLPIS--------SCVTVPD 333


RESULT 6
SNF1_CANAL
ID   SNF1_CANAL     STANDARD;      PRT;   620 AA.
AC   P52497; Q00309;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN   SNF1.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 32354 / B-311;
RX   MEDLINE=98053924; PubMed=9393775;
RA   Petter R., Chang Y.C., Kwon-Chung K.J.;
RT   "A gene homologous to Saccharomyces cerevisiae SNF1 appears to be
RT   essential for the viability of Candida albicans.";
RL   Infect. Immun. 65:4909-4917(1997).
RN   [2]
RP   SEQUENCE OF 7-620 FROM N.A.
RC   STRAIN=ATCC 32354 / B-311;
RA   Petter R., Kwon-Chung K.J.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
CC       INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
CC       PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       SNF1 SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L78129; AAB48643.1; -.
DR   EMBL; L39263; AAA92456.1; -.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Phosphorylation; Carbohydrate metabolism; Nuclear protein.
```

```
FT   DOMAIN       16     29       POLY-HIS.
FT   DOMAIN       53    305       PROTEIN KINASE.
FT   NP_BIND      59     67       ATP (BY SIMILARITY).
FT   BINDING      82     82       ATP (BY SIMILARITY).
FT   ACT_SITE    175    175       BY SIMILARITY.
FT   MOD_RES     208    208       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CONFLICT      7      8       PQ -> AR (IN REF. 2).
FT   CONFLICT    228    234       AGPEVDV -> SSVQKLMI (IN REF. 2).
FT   CONFLICT    242    255       YVMLCGRLPFDDEF -> GMSCCVVDYHSMTSS (IN
FT                                REF. 2).
FT   CONFLICT    387    387       A -> R (IN REF. 2).
FT   CONFLICT    416    416       R -> A (IN REF. 2).
FT   CONFLICT    494    494       S -> L (IN REF. 2).
SQ   SEQUENCE   620 AA;  70005 MW;  1806C652B5061D2B CRC64;

  Query Match             16.1%;  Score 304;  DB 1;  Length 620;
  Best Local Similarity   29.4%;  Pred. No. 2.1e-16;
  Matches   95;  Conservative   55;  Mismatches  131;  Indels   42;  Gaps   12;

Qy         48 LPLSPPTAPDRATAVATASRLGPYVLLEP-EEG--GRAYRALHCPTGTEYTCKVYPVQEA 104
              :|: |   |         |:|:| | :|:   ||  |:   | |  || :   |:   ::
Db         38 VPIDPAANP--------ANRIGRYQILKTLGEGSFGKVKLAQHLGTGQKVALKIIN-RKT 88

Qy        105 LAVLE---------PYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIP 155
              ||  :           | ||  | |: :  :|:      :          ::   :  | ::|
Db         89 LAKSDMQGRVEREISYLRLLRHPHIIKLYDVIKSKDEIIMVIEFAGKELFDYIVQRGKMP 148

Qy        156 EPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDS 215
              | ||   |:|:  |: :||:| :| ||||    :  |:   |:    |  | ::|   |
Db        149 EDEARRFFQQIIAAVEYCHRHKIVHRDLKPENLLLDDQLNVKIADFGL--SNIMT---DG 203

Qy        216 LWDKHAC--PAYV-GPEILSSRASYSGKAADVWSLGVALPTMLAGHYPFQDSEPVLLFGK 272
               : | :|  | |:  ||::| :  |:|   |||| || |: || |  || |     || |
Db        204 NFLKTSCGSPNYMPAPEVISGKL-YAGPEVDVWSAGVILYVMLCGRLPFDDEFIPALFKK 262

Qy        273 IRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEA 332
              |  | | ||  ||| |: |:  :|   |  |:|   |:   | :|| ||
Db        263 ISNGVYTLPNYLSAGAKHLLTRMLVVNPLNRITIHEIMEDDWFKQD-MP---------- 310

Qy        333 AQVVPDGLGLDEAREEEGDREVV 355
                ::|  |  ::  : : | :|:
Db        311 DYLLPPDLSKNKNSKIDVDEDVI 333


RESULT 7
CNK_MOUSE
ID   CNK_MOUSE      STANDARD;      PRT;   631 AA.
AC   Q60806; Q60822; Q9R009;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-
DE   inducible kinase).
GN   CNK OR FNK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=NIH Swiss;
RX   MEDLINE=95247749; PubMed=7730342;
RA   Donohue P.J., Alberts G.F., Guo Y., Winkles J.A.;
RT   "Identification by targeted differential display of an immediate early
RT   gene encoding a putative serine/threonine kinase.";
RL   J. Biol. Chem. 270:10351-10357(1995).
RN   [2]
RP   SEQUENCE OF 333-437 FROM N.A. (ISOFORM 2).
RC   STRAIN=NIH Swiss;
RA   Kauselmann G., Weiler M., Kuhl D.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   FUNCTION, AND PHOSPHORYLATION.
RX   MEDLINE=98343954; PubMed=9677325;
RA   Chase D., Feng Y., Hanshew B., Winkles J.A., Longo D.L., Ferris D.K.;
RT   "Expression and phosphorylation of fibroblast-growth-factor-inducible
RT   kinase (Fnk) during cell-cycle progression.";
RL   Biochem. J. 333:655-660(1998).
CC   -!- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN REGULATING M
CC       PHASE FUNCTIONS DURING THE CELL CYCLE. MAY ALSO BE PART OF THE
CC       SIGNALING NETWORK CONTROLLING CELLULAR ADHESION. IN VITRO, IS ABLE
CC       TO PHOSPHORYLATE CDC25C AND CASEIN (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: BINDS TO THE CALCIUM/INTEGRIN-BINDING PROTEIN (CIB). THIS
CC       INTERACTION PROBABLY OCCURS VIA THE POLO-BOX DOMAIN (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q60806-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q60806-2; Sequence=VSP_004927;
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SKIN.
CC   -!- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
CC       CELLS EXIT MITOSIS.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       CDC5/POLO SUBFAMILY.
CC   -!- SIMILARITY: Contains 2 POLO box domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U21392; AAC52191.1; -.
DR   EMBL; U22434; AAC52192.1; -.
DR   EMBL; AF136586; AAF08369.1; -.
DR   PIR; A57286; A57286.
DR   MGD; MGI:109604; Cnk.
DR   InterPro; IPR000959; POLO_box.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00659; POLO_box; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50078; POLO_BOX; 2.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW   Phosphorylation; Alternative splicing.
FT   DOMAIN       63    315       PROTEIN KINASE.
FT   NP_BIND      69     77       ATP (BY SIMILARITY).
FT   BINDING      92     92       ATP (BY SIMILARITY).
FT   ACT_SITE    186    186       BY SIMILARITY.
FT   DOMAIN      455    518       POLO BOX 1.
FT   DOMAIN      552    622       POLO BOX 2.
FT   VARSPLIC    373    373       L -> LVSGLMRTSIGHPDVRPE (in isoform 2).
FT                                /FTId=VSP_004927.
FT   CONFLICT    386    386       I -> V (IN REF. 2).
SQ   SEQUENCE   631 AA;  70012 MW;  20857341870DB1D2 CRC64;

  Query Match             15.8%;  Score 298.5;  DB 1;  Length 631;
  Best Local Similarity   28.3%;  Pred. No. 5.8e-16;
  Matches   98;  Conservative   48;  Mismatches  133;  Indels   67;  Gaps   14;

Qy         39 PQPRLPPCLLPLSPPTAPDRATAVATAS-RLGPYVLLEPE------------------- 77
              |:|  |   :| :||  |       | || |  | ||  |
Db         10 PRP-FPRAAVPSAPPAGPG---PPANASPRSEPEVLAGPRAPDPPGRLITDPLSGRTYTK 65

Qy         78 -----EGG--RAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLP-----------PHKH 119
```

```
                :||  | | |    :|  |  || |      | :|: |               |:|
Db        66 GRLLGKGGFARCYEATDTESGIAYAVKVIPQSR---VAKPHQREKILNEIELHRDLQHRH 122

Qy       120 VARPTEVLAGTQLLYAFFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGL 178
             : | :        :| |    :   :  : ::|| : |||      ||: : | : || |:
Db       123 IVRFSHHFEDADNIYIFLELCSRKSLAHIWKARHTLLEPEVRYYLRQILSGLKYLHQRGI 182

Qy       179 VLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHAC--PAYVGPEILSSRAS 236
             : ||||  |  |   :| : :   :  |  |:     |  |  | || ||:|  :
Db       183 LHRDLKLGNFFITD--NMELKVGDFGLAARLEPPEQR--KKTICGTPNYVAPEVLLRQG- 237

Qy       237 YSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLL 296
               |  |||||||  ::|:| |  ||: ::    :  |::  | ||| || ||| |:  :|
Db       238 -HGPEADVWSLGCVMYTLLCGSPPFETADLKETYRCIKQVHYTLPASLSLPARQLLAAIL 296

Qy       297 RREPAERLTATGILLHPWLRQ----DPMPLAPTRSHLWEAAQVVPD 338
             |  | :| :   || | :  :    | :|::        :   |||
Db       297 RASPRDRPSIEQILRHDFFTKGYTPDRLPVS--------SCVTVPD 334


RESULT 8
CNK_RAT
ID   CNK_RAT        STANDARD;      PRT;   615 AA.
AC   Q9R011;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-
DE   inducible kinase) (Fragment).
GN   CNK OR FNK.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99452760; PubMed=10523297;
RA   Kauselmann G., Weiler M., Wulff P., Jessberger S., Konietzko U.,
RA   Scafidi J., Staubli U., Bereiter-Hahn J., Strebhardt K., Kuhl D.;
RT   "The polo-like protein kinases Fnk and Snk associate with a Ca(2+)-and
RT   integrin-binding protein and are regulated dynamically with synaptic
RT   plasticity.";
RL   EMBO J. 18:5528-5539(1999).
CC   -!- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN REGULATING M
CC       PHASE FUNCTIONS DURING THE CELL CYCLE. MAY ALSO BE PART OF THE
CC       SIGNALING NETWORK CONTROLLING CELLULAR ADHESION. IN VITRO, IS ABLE
CC       TO PHOSPHORYLATE CDC25C AND CASEIN (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: BINDS TO THE CALCIUM/INTEGRIN-BINDING PROTEIN (CIB). THIS
CC       INTERACTION PROBABLY OCCURS VIA THE POLO-BOX DOMAIN.
CC   -!- SUBCELLULAR LOCATION: WHEN INDUCED, IT TRANSLOCATES INTO THE
CC       DENDRITES OF ACTIVATED NEURONS.
CC   -!- TISSUE SPECIFICITY: CONSTITUVELY EXPRESSED IN POST-MITOTIC
CC       NEURONS.
CC   -!- INDUCTION: BY THE INTENSE ACTIVITY ASSOCIATED WITH SEIZURES.
CC   -!- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
CC       CELLS EXIT MITOSIS (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       CDC5/POLO SUBFAMILY.
CC   -!- SIMILARITY: Contains 2 POLO box domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF136584; AAF08367.1; -.
DR   InterPro; IPR000959; POLO_box.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00659; POLO_box; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50078; POLO_BOX; 2.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW   Phosphorylation.
FT   NON_TER      1      1
FT   DOMAIN      31    283       PROTEIN KINASE.
FT   NP_BIND     37     45       ATP (BY SIMILARITY).
FT   BINDING     60     60       ATP (BY SIMILARITY).
FT   ACT_SITE   154    154       BY SIMILARITY.
FT   DOMAIN     440    503       POLO BOX 1.
FT   DOMAIN     537    607       POLO BOX 2.
FT   NON_TER    615    615
SQ   SEQUENCE   615 AA;  68800 MW;  06584C229B5D71BD CRC64;

  Query Match             15.4%;  Score 291;  DB 1;  Length 615;
  Best Local Similarity   28.5%;  Pred. No. 2.2e-15;
  Matches   94;  Conservative   47;  Mismatches  127;  Indels   62;  Gaps   13;

Qy        41 PRLPPCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEG-------------GRAYRAL 86
             ||  | :  |: |  ||       ||||    : :|  |                | | |
Db         3 PRSEPEV--LAGPQTPD-------ASRL----ITDPRSGRTYIKGRLLGKGGFARCYEAT 49

Qy        87 HCPTGTEYTCKVYPVQEALAVLEPYARLP----------PHKHVARPTEVLAGTQLLYA 135
               |   | || |      | :|: |                |:|: | :       :|
Db        50 DTETSIAYAVKVIPQSR---VAKPHQREKIINEIELHRDLQHRHIVRFSHHFEDADNIYI 106

Qy       136 FFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRE 194
             |    :   :  : ::|| : |||      ||: : | : || |:: ||||  |   |
Db       107 FLELCSRKSLAHIWKARHTLLEPEVRYYLRQILSGLKYLHQRGILHRDLKLGNFFITD-- 164

Qy       195 RKKLVLENLEDSCVLTGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALF 252
               :| : :   :  |  |:     |  |  | || ||:|  :    |  |||||||  ::
Db       165 NMELKVGDFGLAARLEPPEQR--KKTICGTPNYVAPEVLLRQG--HGPEADVWSLGCVMY 220

Qy       253 TMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLH 312
             |:| |  ||: ::    :  |::  | ||| || ||| |:  :||  | :| :   || |
Db       221 TLLCGSPPFETADLKETYRCIKQVHYTLPASLSLPARQLLAAILRASPRDRPSIEQILRH 280

Qy       313 PWLRQ----DPMPLAPTRSHLWEAAQVVPD 338
               :  :    | :|::        :   |||
Db       281 DPFTKGYTPDRLPVS--------SCVTVPD 302


RESULT 9
KKK1_YEAST
ID   KKK1_YEAST     STANDARD;      PRT;  1518 AA.
AC   P34244;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.-).
GN   YKL101W OR YKL453.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94078677; PubMed=8256524;
RA   Dallier C., Valens M., Puzos V., Pukuhara H., Cheret G., Sor F.,
RA   Bolotin-Fukuhara M.;
RT   "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
RT   physically localizes the MRB1 gene and reveals eight new open reading
RT   frames, including a homologue of the KIN1/KIN2 and SNF1 protein
RT   kinases.";
RL   Yeast 9:1149-1155(1993).
```

```
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       NIM1 SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X71133; CAA50456.1; -.
DR   EMBL; Z28101; CAA81941.1; -.
DR   PIR; S37928; S37928.
DR   HSSP; Q63450; 1A06.
DR   SGD; S0001584; HSL1.
DR   GO; GO:0005935; C:bud neck; IDA.
DR   GO; GO:0005940; C:septin ring; IDA.
DR   GO; GO:0000086; P:G2/M transition of mitotic cell cycle; IGI.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR   GO; GO:0000074; P:regulation of cell cycle; IMP.
DR   GO; GO:0000135; P:septin checkpoint; IGI.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW   ATP-binding.
FT   DOMAIN       81    369       PROTEIN KINASE.
FT   NP_BIND      87     95       ATP (BY SIMILARITY).
FT   BINDING     110    110       ATP (BY SIMILARITY).
FT   ACT_SITE    239    239       BY SIMILARITY.
SQ   SEQUENCE   1518 AA;  169592 MW;  803F84F7531241DD CRC64;

  Query Match             15.2%;  Score 287;  DB 1;  Length 1518;
  Best Local Similarity   26.7%;  Pred. No. 1.3e-14;
  Matches   86;  Conservative   50;  Mismatches  128;  Indels   58;  Gaps    9;

Qy          56 PDRATAVATAS-------RLGPYVL---LEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
               ||   :||| |         :||: |   |     ||   | :  ||      |: | ::|
Db          59 PDSTVSVATKSSKRKSRDTVGPWKLGKTLGKGSSGRVRLAKNMETGQLAAIKIVPKKKAF 118

Qy         106 ----------------------------------AVLEPYA--------RLPPHKHVAR 122
                                               :   ||          :|  | :|
Db         119 VHCSNNGTVPNSYSSSMVTSNVSSPSIASREHSNHSQTNPYGIEREIVIMKLISHTNVMA 178

Qy         123 PTEVLAGTQLLYAFFTRTH-GDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLR 181
               ||      ||       |::   : |: ::|| ||    |:|:   :::||    :  |
Db         179 LFEVWENKSELYLVLEYVDGGELFDYLVSKGKLPEREAIHYFKQIVEGVSYCHSFNICHR 238

Qy         182 DLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKA 241
               |||     :  |:: ::: : :       |  |:  |      | |  |||:  |  | |
Db         239 DLK-PENLLLDKKNRRIKIADF-GMAALELPNKLLKTSCGSPHYASPEIVMGR-PYHGGP 295

Qy         242 ADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPA 301
               :|||| |: || :| || || |      |  |:: | | :|: ||: || |:  :|  :|
Db         296 SDVWSCGIVLFALLTGHLPFNDDNIKKLLLKVQSGKYQMPSNLSSEARDLISKILVIDPE 355

Qy         302 ERLTATGILLHPWLRQ-DPMPL 322
               :|:|   || || ::: | :|:
Db         356 KRITTQEILKHPLIKKYDDLPV 377


RESULT 10
SNF1_CANGA
ID   SNF1_CANGA     STANDARD;      PRT;   611 AA.
AC   Q00372;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN   SNF1.
OS   Candida glabrata (Yeast) (Torulopsis glabrata).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5478;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCCL884;
RX   MEDLINE=97101049; PubMed=8945576;
RA   Petter R., Kwon-Chung K.J.;
RT   "Disruption of the SNF1 gene abolishes trehalose utilization in the
RT   pathogenic yeast Candida glabrata.";
RL   Infect. Immun. 64:5269-5273(1996).
CC   -!- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
CC       INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
CC       PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       SNF1 SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L78130; AAB48642.1; -.
DR   HSSP; P24941; 1HCL.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Phosphorylation; Carbohydrate metabolism; Nuclear protein.
FT   DOMAIN       6     17       POLY-HIS.
FT   DOMAIN      39    290       PROTEIN KINASE.
FT   NP_BIND     45     53       ATP (BY SIMILARITY).
FT   BINDING     68     68       ATP (BY SIMILARITY).
FT   ACT_SITE   161    161       BY SIMILARITY.
FT   MOD_RES    194    194       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ   SEQUENCE   611 AA;  70049 MW;  89E17812A4900CD0 CRC64;

  Query Match             15.0%;  Score 284.5;  DB 1;  Length 611;
  Best Local Similarity   27.9%;  Pred. No. 7e-15;
  Matches   90;  Conservative   54;  Mismatches  136;  Indels   43;  Gaps   10;

Qy          57 DRATAVATASRLGPYVLLEP-EEG--GRAYRALHCPTGTEYTCKVYPVQEALAVLE---- 109
               :: :::|  ||:| | :::   ||  |:   | |  || :   |    :: ||  :
Db          25 NKVSSLADGSRVGNYQIVKTLGEGSFGKVKLAYHVTTGQKVALKSIN-KKVLAKSDMQGR 83

Qy         110 -----PYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFR 164
                    | ||  | |: :  :|:     :           ::   :  |::: | ||    |:
Db          84 IDREISYLRLLRHPHIIKLYDVIKSKDEIIMVIEYAGNELFDYIVQRNKMSEQEARRFFQ 143

Qy         165 QMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHAC-- 222
               |: :|: :||:| :| ||||    :  :    |:    |  | ::|   |  : | :|
Db         144 QIISAVEYCHRHKIVHRDLKPENLLLDEHLNVKIADFGL--SNIMT---DGNFLKTSCGS 198

Qy         223 PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPA 282
               | |  ||::| :  |:|   |||| || |: ||     || |    :||  |  | | ||
Db         199 PNYAAPEVISGKL-YAGPEVDVWSCGVILYVMLCRRLPFDDESIPVLFKNISNGVYTLPK 257
```

```
Qy        283 GLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGL-- 340
Db        258 FLSPGASDLIKRMLIVNPLNRISIHEIMQDEWFKVD-----------LAEYLVPQDLKQ 305

Qy        341 --------GLDEAREEEGDREVV 355
Db        306 QEQFNKKKSGNEENVEEIDDEMVV 328
```

```
RESULT 11
KI10_ARATH
ID   KI10_ARATH     STANDARD;      PRT;   512 AA.
AC   Q38997; O04728; Q39076;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   SNF1-related protein kinase KIN10 (EC 2.7.1.-) (AKIN10).
GN   KIN10 OR SKIN10 OR AT3G01090 OR T4P13.22.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93013041; PubMed=1339373;
RA   le Guen L., Thomas M., Bianchi M., Halford N.G., Kreis M.;
RT   "Structure and expression of a gene from Arabidopsis thaliana
RT   encoding a protein related to SNF1 protein kinase.";
RL   Gene 120:249-254(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lessard P., Kreis M., Thomas M.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016720; PubMed=11130713;
RA   Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA   Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA   Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA   De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA   Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA   Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA   Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA   Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA   Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA   Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA   Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA   Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA   Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA   de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA   Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA   Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA   Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA   Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA   Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA   Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA   Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA   Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA   Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT   "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:820-822(2000).
RN   [4]
RP   SEQUENCE OF 1-19 FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=95115691; PubMed=7816049;
RA   le Guen L., Thomas M., Kreis M.;
RT   "Gene density and organization in a small region of the Arabidopsis
RT   thaliana genome.";
RL   Mol. Gen. Genet. 245:390-396(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN A SIGNAL TRANSDUCTION
CC       CASCADE REGULATING GENE EXPRESSION AND CARBOHYDRATE METABOLISM IN
CC       HIGHER PLANTS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ROOTS, SHOOTS AND LEAVES.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       SNF1 SUBFAMILY.
CC   -!- SIMILARITY: Contains 1 UBA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M93023; AAA32736.1; -.
DR   EMBL; X79707; CAA56146.1; -.
DR   EMBL; AC008261; AAF26165.1; -.
DR   EMBL; X94757; CAA64384.1; -.
DR   PIR; JC1446; JC1446.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR001772; Kinase_Cterm.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   InterPro; IPR000449; UBA_domain.
DR   Pfam; PF02149; KA1; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00627; UBA; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00165; UBA; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50030; UBA; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Phosphorylation.
FT   DOMAIN       19    271       PROTEIN KINASE.
FT   DOMAIN      292    332       UBA.
FT   NP_BIND      25     33       ATP (BY SIMILARITY).
FT   BINDING      48     48       ATP (BY SIMILARITY).
FT   ACT_SITE    142    142       BY SIMILARITY.
FT   MOD_RES     175    175       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ   SEQUENCE   512 AA;  58373 MW;  5A18655A0AA506DF CRC64;

  Query Match             14.9%;  Score 281;  DB 1;  Length 512;
  Best Local Similarity   28.7%;  Pred. No. 1.1e-14;
  Matches   87;  Conservative   46;  Mismatches  118;  Indels   52;  Gaps    9;

Qy         80 GRAYRALHCPTGTEYTCKVY-----------PVQEALAVLEPYARLPPHKHVARPTEVL 127
Db         31 GRVKIAEHALTGHKVAIKILNRRKIKNMEMEEKVRREIKIL----RLFMHPHIIRLYEVI 86

Qy        128 AGTQLLYAFFTRTH-GDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLK-- 184
Db         87 ETPTDIYLVMEYVNSGELFDYIVEKGRLQEDEARNFFQQIISGVEYCHRNMVVHRDLKPE 146

Qy        185 ------LCRFVFADRERKKLVLEN--LEDSCVLTGPDDSLWDKHACPAYVGPEILSSRAS 236
Db        147 NLLLDSKCNVKIADFGLSNIMRDGHFLKTSC------------GSPNYAAPEVISGKL- 192

Qy        237 YSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLL 296
Db        193 YAGPEVDVWSCGVILYALLCGTLPFDDENIPNLFKKIKGGIYTLPSHLSPGARDLIPRML 252

Qy        297 RREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQV-VPDGLGLDEAREEEGDREVV 355
```

```
Db        253 VVDPMKRVTIPEIRQHPWF----------QAHLPRYLAVPPPDTVQQAKKIDEEILQEVI 302

Qy        356 LYG 358
                |
Db        303 NMG 305


RESULT 12
SNF1_YEAST
ID   SNF1_YEAST     STANDARD;      PRT;   633 AA.
AC   P06782;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN   SNF1 OR CAT1 OR CCR1 OR PAS14 OR GLC2 OR YDR477W OR D8035.20.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86289463; PubMed=3526554;
RA   Celenza J.L., Carlson M.;
RT   "A yeast gene that is essential for release from glucose repression
RT   encodes a protein kinase.";
RL   Science 233:1175-1180(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
RA   Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA   Hunicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA   Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA   Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA   Winant A., Yelton M., Botstein D., Davis R.W.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 274-284; 528-539 AND 622-630, AND PHOSPHORYLATION SITE.
RX   MEDLINE=94131988; PubMed=7905477;
RA   Mitchelhill K.I., Stapleton D., Gao G., House C., Michell B.,
RA   Katsis F., Witters L.A., Kemp B.E.;
RT   "Mammalian AMP-activated protein kinase shares structural and
RT   functional homology with the catalytic domain of yeast Snf1 protein
RT   kinase.";
RL   J. Biol. Chem. 269:2361-2364(1994).
CC   -!- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
CC       INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
CC       PROTEIN SNF4. INTERACTS ALSO WITH SIP1, SIP2 AND GAL83. COULD
CC       PHOSPHORYLATES CAT8.
CC   -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       SNF1 SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M13971; AAA35058.1; -.
DR   EMBL; U33050; AAB64904.1; -.
DR   PIR; A26030; A26030.
DR   HSSP; P24941; 1HCL.
DR   SGD; S0002885; SNF1.
DR   GO; GO:0005737; C:cytoplasm; IPI.
DR   GO; GO:0005634; C:nucleus; IPI.
DR   GO; GO:0000324; C:vacuole (sensu Fungi); IPI.
DR   GO; GO:0004679; F:SNF1A/AMP-activated protein kinase activity; IDA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR   GO; GO:0006995; P:response to nitrogen starvation; IDA.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Phosphorylation; Carbohydrate metabolism; Nuclear protein.
FT   DOMAIN       18     32       POLY-HIS.
FT   DOMAIN       55    306       PROTEIN KINASE.
FT   NP_BIND      61     69       ATP (BY SIMILARITY).
FT   BINDING      84     84       ATP (BY SIMILARITY).
FT   ACT_SITE    177    177       BY SIMILARITY.
FT   MOD_RES     210    210       PHOSPHORYLATION (AUTO-).
SQ   SEQUENCE   633 AA;  72045 MW;  F5C63565C986C4B3 CRC64;

  Query Match             14.5%;  Score 274.5;  DB 1;  Length 633;
  Best Local Similarity   27.8%;  Pred. No. 4.4e-14;
  Matches   88;  Conservative   53;  Mismatches  138;  Indels   37;  Gaps   11;

Qy         51 SPPTAPDRATAVATASRLGPYVLLEP-EEG--GRAYRALHCPTGTEYTCKVYPVQEALAV 107
              |  |   :  :::|  : :| | :::   ||  |:   | |  || :   |:   :: ||
Db         35 SNSTLNNPKSSLADGAHIGNYQIVKTLGEGSFGKVKLAYHTTTGQKVALKIIN-KKVLAK 93

Qy        108 LE---------PYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPE 158
               :           | ||  | |: :  :|:     :          ::   :  | :: | |
Db         94 SDMQGRIERESISYLRLLRHPHIIKLYDVIKSKDEIIMVIEYAGNELFDYIVQRDKMSEQE 153

Qy        159 AAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD 218
              |   |:|: :|: :||:| :| ||||    :  :    |:    |  | ::|   |  :
Db        154 ARRFFQQIISAVEYCHRHKIVHRDLKPENLLLDEHLNVKIADFGL--SNIMT---DGNFL 208

Qy        219 KHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG 276
              | :|  | |  ||::| :  |:|   |||| || |: ||    || |    :|| |  |
Db        209 KTSCGSPNYAAPEVISGKL-YAGPEVDVWSCGVILYVMLCRRLPFDDESIPVLFKNISNG 267

Qy        277 AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD-PMPLAPTRSHLWEAAQV 335
              | ||  ||  |  |:: :|   |  |::   |:   | : | |  | |
Db        268 VYTLPKFLSPGAAGLIKRMLIVNPLNRISIHEIMQDDWFKVDLPEYLLP---------- 316

Qy        336 VPDGLGLDEAREEEGD 351
               ||   |    ||| :
Db        317 -PD---LKPHPEEENE 328


RESULT 13
KCCD_RAT
ID   KCCD_RAT       STANDARD;      PRT;   533 AA.
AC   P15791;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Calcium/calmodulin-dependent protein kinase type II delta chain
DE   (EC 2.7.1.123) (CaM-kinase II delta chain) (CaM kinase II delta
DE   subunit) (CaMK-II delta subunit).
GN   CAMK2D.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90036861; PubMed=2553697;
RA   Tobimatsu T., Fujisawa H.;
RT   "Tissue-specific expression of four types of rat calmodulin-dependent
RT   protein kinase II mRNAs.";
RL   J. Biol. Chem. 264:17907-17912(1989).
RN   [2]
RP   SEQUENCE OF 314-533 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Aorta, and Skeletal muscle;
RX   MEDLINE=93300844; PubMed=8390994;
RA   Schworer C.M., Rothblum L.I., Thekkumkara T.J., Singer H.A.;
```

```
RT   "Identification of novel isoforms of the delta subunit of
RT   Ca2+/calmodulin-dependent protein kinase II. Differential expression
RT   in rat brain and aorta.";
RL   J. Biol. Chem. 268:14443-14449(1993).
CC   -!- FUNCTION: THIS KINASE MAY PLAY A ROLE IN NEUROTRANSMISSION.
CC   -!- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.
CC   -!- ENZYME REGULATION: AUTOPHOSPHORYLATION OF CAM-KINASE II PLAYS AN
CC       IMPORTANT ROLE IN THE REGULATION OF THE KINASE ACTIVITY.
CC   -!- SUBUNIT: COMPOSED OF FOUR DIFFERENT CHAINS: ALPHA, BETA, GAMMA,
CC       AND DELTA.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=Delta 1;
CC         IsoId=P15791-1; Sequence=Displayed;
CC       Name=Delta 2;
CC         IsoId=P15791-2; Sequence=VSP_004784;
CC       Name=Delta 3;
CC         IsoId=P15791-3; Sequence=VSP_004785, VSP_004786, VSP_004787;
CC       Name=Delta 4;
CC         IsoId=P15791-4; Sequence=VSP_004788;
CC   -!- TISSUE SPECIFICITY: DELTA 1 IS THE PREDOMINANT FORM IN THE BRAIN,
CC       DELTA 2 AND 3 PREDOMINATE IN THE AORTA AND DELTA 4 IN SKELETAL
CC       MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. CAMK
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J05072; AAA40866.1; -.
DR   EMBL; L13406; AAA41479.1; -.
DR   EMBL; L13407; AAA41480.1; -.
DR   EMBL; L13408; AAA41481.1; -.
DR   PIR; A34366; A34366.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW   Phosphorylation; ATP-binding; Neurone; Alternative splicing.
FT   DOMAIN       14    272       PROTEIN KINASE.
FT   NP_BIND      20     28       ATP (BY SIMILARITY).
FT   BINDING      43     43       ATP (BY SIMILARITY).
FT   ACT_SITE    136    136       BY SIMILARITY.
FT   DOMAIN      291    301       CALMODULIN-BINDING (BY SIMILARITY).
FT   VARSPLIC    329    362       Missing (in isoform Delta 2).
FT                                /FTId=VSP_004784.
FT   VARSPLIC    329    335       INNKANV -> KRKSSSV (in isoform Delta 3).
FT                                /FTId=VSP_004785.
FT   VARSPLIC    337    359       Missing (In isoform Delta 3).
FT                                /FTId=VSP_004786.
FT   VARSPLIC    360    362       GNK -> QMM (in isoform Delta 3).
FT                                /FTId=VSP_004787.
FT   VARSPLIC    349    362       Missing (In isoform Delta 4).
FT                                /FTId=VSP_004788.
SQ   SEQUENCE   533 AA;  60080 MW;  E41BCB2B5A00E7CA CRC64;

  Query Match             14.5%;  Score 273.5;  DB 1;  Length 533;
  Best Local Similarity   28.8%;  Pred. No. 4.3e-14;
  Matches   79;  Conservative   53;  Mismatches  123;  Indels   19;  Gaps    7;

Qy         61 AVATASRLGPYVLLEPEEGGRAY----RALHCPTGTEYTCKVYPVQEALA----VLEPYA 112
              :  | :|      |  | |  |:    | :  ||| ||  |:   ::  |     ||  |

Db          3 STTTCTRFTDEYQLFEELGKGAFSVVRRCMKIPTGQEYAAKIINTKKLSARDHQKLERBA 62

Qy        113 ---RLPPHKHVARPTEVLAGTQLLYAPFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMAT 168
                 ||  | :: |  : ::     |  |    | |::   : :|    | :|:   :|:
Db         63 RICRLLKHPNIVRLHDSISEEGFHYLVFDLVTGGELFEDIVARBYYSEADASHCIQQILE 122

Qy        169 ALAHCHQHGLVLRDLKLCRFVFADRERKKLV-LENLEDSCVLTGPDDSLWDKHACPAYVG 227
              :: ||| :|:| ||||    : | : :   | | :   :  : |   : :      | |:
Db        123 SVNHCHLNGIVHRDLKPENLLLASKSKGAAVKLADFGLAIEVQGDQQAWFGFAGTPGYLS 182

Qy        228 PEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPA----G 283
              ||:|  |    ||  |:|: || |: :| |: || | :   |: :|: |||   |:
Db        183 PEVL--RKDPYGKPVDMWACGVILYILLVGYPPFWDEDQHRLYQQIKAGAYDFPSPEWDT 240

Qy        284 LSAPARCLVRCLLRREPAERLTATGILLHPWLRQ 317
              ::  |: |:  :|   ||:|:||:  | |||: |
Db        241 VTPEAKDLINKMLTINPAKRITASEALKHPWICQ 274


RESULT 14
CDR1_SCHPO
ID   CDR1_SCHPO     STANDARD;      PRT;   593 AA.
AC   P07334; Q9P6Q4;
DT   01-APR-1988 (Rel. 07, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Mitosis inducer protein kinase cdr1 (EC 2.7.1.-) (Protein kinase
DE   nim1).
GN   CDR1 OR NIM1 OR SPAC644.06C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91169281; PubMed=2004705;
RA   Feilotter H., Nurse P., Young P.G.;
RT   "Genetic and molecular analysis of cdr1/nim1 in Schizosaccharomyces
RT   pombe.";
RL   Genetics 127:309-318(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87187654; PubMed=3453113;
RA   Russell P., Nurse P.;
RT   "The mitotic inducer nim1+ functions in a regulatory network of
RT   protein kinase homologs controlling the initiation of mitosis.";
RL   Cell 49:569-576(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
```

```
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- FUNCTION: THIS PROTEIN, A DOSE-DEPENDENT MITOTIC INDUCER, APPEARS
CC       TO FUNCTION AS A NEGATIVE REGULATOR OF MITOSIS INHIBITOR WEE1 BY
CC       PHOSPHORYLATING AND INACTIVATING IT.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       NIM1 SUBFAMILY.
CC   -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
CC       FRAMESHIFTS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X57549; CAA40774.1; -.
DR   EMBL; M16509; AAA35317.1; ALT_FRAME.
DR   EMBL; AL355012; CAB90133.1; -.
DR   PIR; S16153; KIZPMN.
DR   HSSP; Q63450; 1A06.
DR   GeneDB_SPombe; SPAC644.06c; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Mitosis.
FT   DOMAIN       12    258       PROTEIN KINASE.
FT   NP_BIND      18     26       ATP (BY SIMILARITY).
FT   BINDING      41     41       ATP (BY SIMILARITY).
FT   ACT_SITE    128    128       BY SIMILARITY.
FT   CONFLICT    252    252       V -> P (IN REF. 1).
FT   CONFLICT    570    570       A -> I (IN REF. 1).
SQ   SEQUENCE   593 AA;  66954 MW;  DE30AE06B070F458 CRC64;

  Query Match             14.4%;  Score 271.5;  DB 1;  Length 593;
  Best Local Similarity   31.5%;  Pred. No. 7e-14;
  Matches   74;  Conservative   45;  Mismatches  107;  Indels    9;  Gaps    6;

Qy         85 ALHCPTGTEYTCKVYPVQEALAVLE-PYARLPPHKHVARPTEVLAGTQLLY-AFFTRTHG 142
              | |  ||     |: |:: |   :|    ||  | :: |  :|    | :| |      |
Db         29 AKHAKTGDLAAIKIIPIRYASIGMEILMMRLLRHPNILRLYDVWTDHQHMYLALEYVPDG 88

Qy        143 DMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLEN 202
             ::   :|    : | |||    |: |:||||:     |||||   :    |::   ::
Db         89 ELFHYIRKHGPLSEREAAHYLSQILDAVAHCHRFRVFHRDLKLENILIKVNEQQ---IKI 145

Qy        203 LEDSCVLTGPDDSLWDKHACPA--YVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYP 260
              :      |:||  : : | :  |: |||:| :  | |  ||||| || |:::|:   |
Db        146 ADFGMATVEPNDSCLENY-CGSLHYLAPEIVSHK-PYRGAPADVWSCGVILYSLLSNKLP 203

Qy        261 FQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWL 315
              |.    ::: ||| ||| ||: :|: |: |:  :|   |: |:|   :  ||:|
Db        204 FGGQNTDVIYNKIRHGAYDLPSSISSAAQDLLHRMLDVNPSTRITIPEVFSHPFL 258


RESULT 15
SNF1_SCHPO
ID   SNF1_SCHPO     STANDARD;      PRT;   576 AA.
AC   O74536;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   SNF1-like protein kinase (EC 2.7.1.-).
GN   SPCC74.03C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       SNF1 SUBFAMILY.
CC   -!- SIMILARITY: Contains 1 UBA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL031543; CAA20833.1; -.
DR   PIR; T41587; T41587.
DR   HSSP; Q63450; 1A06.
DR   GeneDB_SPombe; SPCC74.03c; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR000449; UBA_domain.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50030; UBA; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding.
FT   DOMAIN       34    285       PROTEIN KINASE.
FT   DOMAIN      304    345       UBA.
FT   NP_BIND      40     48       ATP (BY SIMILARITY).
FT   BINDING      63     63       ATP (BY SIMILARITY).
FT   ACT_SITE    156    156       BY SIMILARITY.
SQ   SEQUENCE   576 AA;  65996 MW;  E5857E8F171E7B50 CRC64;
```

```
  Query Match             14.3%;  Score 271;  DB 1;  Length 576;
  Best Local Similarity   25.8%;  Pred. No. 7.5e-14;
  Matches   83;  Conservative   54;  Mismatches  133;  Indels   52;  Gaps   11;

Qy          52 PPTAPDRATAVATASRLGPYVLLEP-EEG--GRAYRALHCPTGTEYTCKVYP-------- 100
               || |         :   :|||:: |   ||  |:   | |  |  :   |
Db          22 PPEA-------ISKRHIGPYIIRETLGEGSFGKVKLATHYKTQQKVALKFISRQLLKKSD 74

Qy         101 ----VQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPE 156
                 |:  ::    | :|  | |: :  :|:     :          |::   :  : |: |
Db          75 MHMRVEREIS----YLKLLRHPHIIKLYDVITTPTDIVMVIEYAGGELFDYIVEKKRMTE 130

Qy         157 PEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSL 216
                |    |:|:  |: :||:| :| ||||    :  |    |:    |  | ::|   |
Db         131 DEGRRFFQQIICAIEYCHRHKIVHRDLKPENLLLDDNLNVKIADFGL--SNIMT---DGN 185

Qy         217 WDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIR 274
               : | :|  | |  ||::: :  |:|   |||| |: |: || |  || |     || |:
Db         186 FLKTSCGSPNYAAPEVINGKL-YAGPEVDVWSCGIVLYVMLVGRLPFDDEFIPNLFKKVN 244

Qy         275 RGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ 334
                 | :|  ||  |: |:| ::  :| :|:|   |   ||   :
Db         245 SCVYVMPDFLSPGAQSLIRRMIVADPMQRITIQEIRRDPWFNVN---------------- 288

Qy         335 VVPDGL-GLDEAREEEGDREVV 355
                :|| |  ::| :    |  :|
Db         289 -LPDYLRPMEEVQGSYADSRIV 309


Search completed: January 15, 2004, 14:56:19
Job time : 19 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:55:46 ; Search time 20 Seconds
(without alignments)
1721.419 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 1891
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	298.5	15.8	631	2	A57286	probable serine/th
2	298	15.8	887	2	T20941	hypothetical prote
3	292.5	15.5	651	2	S52244	p69Eg3 protein - A
4	291	15.4	602	2	S72513	FOG2 protein - yea
5	287	15.2	1518	2	S37928	probable purine nu
6	284	15.0	504	2	T10449	probable serine/th
7	284	15.0	512	2	T52633	serine/threonine-s
8	281	14.9	512	1	JC1446	serine/threonine-s
9	280	14.8	339	2	S56719	serine/threonine-s
10	277.5	14.7	502	2	T02306	probable protein k
11	277	14.6	441	2	E85362	hypothetical prote
12	277	14.6	511	1	A56009	serine/threonine-s
13	274.5	14.5	633	1	A26030	serine/threonine-s
14	274	14.5	513	1	S60304	serine/threonine-s
15	273.5	14.5	533	1	A34366	Ca2+/calmodulin-de
16	273	14.4	504	2	T07415	probable serine/th
17	272.5	14.4	480	2	A86427	probable serine/th
18	271.5	14.4	421	2	E96522	hypothetical prote
19	271	14.3	431	2	T02496	probable protein k
20	271	14.3	576	2	T41587	probable carbon ca
21	270.5	14.3	426	2	C71408	probable protein k
22	270	14.3	512	2	T07788	probable serine/th
23	269	14.2	542	1	A45025	Ca2+/calmodulin-de
24	268.5	14.2	593	1	KIZPMN	protein kinase cdr
25	268	14.2	622	1	S44859	serine/threonine-s
26	267	14.1	472	2	B90100	SNF-related kinase
27	266	14.1	542	1	A26464	Ca2+/calmodulin-de
28	266	14.1	589	2	S68470	Ca2+/calmodulin-de
29	264	14.0	442	2	T48203	hypothetical prote
30	263	13.9	1142	2	S59359	GIN4 protein - yea
31	261.5	13.8	421	2	T48202	protein kinase AK2
32	261.5	13.8	513	1	S60303	serine/threonine-s
33	261.5	13.8	552	1	S51025	[hydroxymethylglut
34	261.5	13.8	891	2	T40503	protein kinase kin
35	260.5	13.8	473	1	S59941	serine/threonine-s
36	260.5	13.8	552	1	A53621	[hydroxymethylglut
37	260.5	13.8	891	2	A38903	protein kinase 1 -
38	259	13.7	350	2	T06107	probable serine/th
39	258.5	13.7	440	2	T14736	probable serine/th
40	258.5	13.7	445	2	T50802	serine/threonine p
41	258	13.6	502	1	A41361	serine/threonine-s
42	258	13.6	726	2	T33998	hypothetical prote
43	257.5	13.6	492	1	T03271	calcium-dependent
44	257	13.6	469	2	B84644	probable protein k
45	256.5	13.6	516	1	JU0270	Ca2+/calmodulin-de

ALIGNMENTS

RESULT 1
A57286
probable serine/threonine protein kinase (EC 2.7.1.-) fnk - mouse
C;Species: Mus musculus (house mouse)
C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 10-Sep-1997
C;Accession: A57286
R;Donohue, P.J.; Alberts, G.F.; Guo, Y.; Winkles, J.A.
J. Biol. Chem. 270, 10351-10357, 1995
A;Title: Identification by targeted differential display of an immediate early gene en
A;Reference number: A57286; MUID:95247749; PMID:7730342
A;Accession: A57286
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-631 <DON>
A;Cross-references: GB:U21392; GB:U22434
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homc
C;Keywords: ATP; phosphotransferase
F;61-315/Domain: protein kinase homology <KIN>

Query Match 15.8%; Score 298.5; DB 2; Length 631;
Best Local Similarity 28.3%; Pred. No. 4.5e-13;
Matches 98; Conservative 48; Mismatches 133; Indels 67; Gaps 14;

```
Qy         39 PQPRLPPCLLPLSPPTAPDRATAVATAS-RLGPYVLLEPE---------------------- 77
              |:|  |   :| :||  |       | || |  | ||  |
Db         10 PRP-PPRAAVPSAPPAGPG---PPANASPRSEPEVLAGPRAPDPPGRLITDPLSGRTYTK 65

Qy         78 -----EGG--RAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLP-----------PHKH 119
                  :||  | | |    :|  |  || |       | :|: |              |:|
Db         66 GRLLGKGGFARCYEATDTESGIAYAVKVIPQSR---VAKPHQREKILNEIELHRDLQHRH 122

Qy        120 VARPTEVLAGTQLLYAPFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGL 178
             : | :     :| |    :  :  : ::|| : |||     ||: : | : || |:
Db        123 IVRFSHHFEDADNIYIFLELCSRKSLAHIWKARHTLLEPEVRYYLRQILSGLKYLHQRGI 182

Qy        179 VLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHAC--PAYVGPEILSSRAS 236
             : ||||| |   |    :| : :  :  | |:     |  |  | || ||:|  :
Db        183 LHRDLKLGNFFITD--NMELKVGDFGLAARLEPPEQR--KKTICGTPNYVAPEVLLRQG- 237

Qy        237 YSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLL 296
               | |||||||  ::|:| |  ||: ::    :  |::  | ||| || ||| |:  :|
Db        238 -HGPEADVWSLGCVMYTLLCGSPPFETADLKETYRCIKQVHYTLPASLSLPARQLLAAIL 296

Qy        297 RREPAERLTATGILLHPWLRQ----DPMPLAPTRSHLWEAAQVVPD 338
              |  | :| :   || | :  :    | :|::        :  |||
Db        297 RASPRDRPSIEQILRHDFFTKGYTPDRLPVS--------SCVTVPD 334
```

RESULT 2
T20941
hypothetical protein F15A2.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans

```
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T20941
R;Gregory, J.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19349
A;Accession: T20941
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-887 <WIL>
A;Cross-references: EMBL:Z70207; PIDN:CAA94127.1; GSPDB:GN00028; CESP:F15A2.6
A;Experimental source: clone F15A2
C;Genetics:
A;Gene: CESP:F15A2.6
A;Map position: X
A;Introns: 32/1; 63/3; 92/2; 139/2; 189/3; 328/2; 448/2; 516/3; 604/2; 684/3; 735/2; 777

  Query Match             15.8%;  Score 298;  DB 2;  Length 887;
  Best Local Similarity   25.9%;  Pred. No. 7.1e-13;
  Matches   95;  Conservative   51;  Mismatches  133;  Indels   88;  Gaps   14;

Qy         62 VATASRLGPYVL---LEPEEGGRAYRALHCPTGTEYTCKV-----------YPVQEALAV 107
              || |   ||| |    |   : |       || || :   |:             |:  :|:
Db         11 VAQAQYCGPYKLEKTLGKGQTGLVKTGTHCITGRKVAIKIVNKEKLSESVLQKVEREIAI 70

Qy        108 LEPYARLPPHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVRSRHRIPEPEAAVLFRQM 166
              :    :|  | ||    :|    : ||      : |::  :  : |:   ||   |||:
Db         71 M----KLIEHPHVLHLYDVYENKKYLYLLLEHVSGGELFDYLVRKGRLMSKEARKFFRQI 126

Qy        167 ATALAHCHQHGLVLRDLKLCRFVFADRERKK--------LVLEN--LEDSCVLTGPDDSL 216
               :||  || | :  ||||    :  :|   |         | :|   || ||
Db        127 ISALDFCHAHNICHRDLKPENLLLDERNNIKVADFGMASLQVEGSMLETSC--------- 177

Qy        217 WDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG 276
                  | |  ||::      | |: ||||| || |: :| |  || |      |  |::||
Db        178 ----GSPHYACPEVIRGE-KYDGRKADVWSCGVILYALLVGALPFDDDNLRNLLEKVKRG 232

Qy        277 AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWL----RQDP---MPLAP-TRSH 328
               : :|  : |  : |:| ::  :| :| :   :  |||:    : ||   :|::   ::|
Db        233 VFHIPHFVPADVQSLLRAMIEVDPGKRYSLADVFKHPWVSGTTKADPELELPMSQVVQTH 292

Qy        329 LWEAAQVV-PDGLG-------------------------------LDEARE---EEGD 351
              :      : || |                               ||  |     :| |
Db        293 VIPGEDSIDPDVLRHMNCLGCFKDKQKLINELLSPKHNTEKMVYFLLLDRKRRRPAQEDD 352

Qy        352 REVVLYG 358
               |:|| |
Db        353 TEIVLRG 359


RESULT 3
S52244
p69Eg3 protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 07-May-1995 #sequence_revision 03-Aug-1995 #text_change 24-May-2001
C;Accession: S52244
R;Roghi, C.; le Guellec, R.; Paris, J.; Couturier, A.; Philippe, M.
submitted to the EMBL Data Library, October 1992
A;Description: Eg3, selected by differential screening encodes a new Xenopus protein kin
A;Reference number: S52243
A;Accession: S52244
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-651 <ROG>
A;Cross-references: EMBL:Z17205; NID:g609283; PIDN:CAA78913.1; PID:g609284
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP
F;11-265/Domain: protein kinase homology <KIN>
F;19-27/Region: protein kinase ATP-binding motif

  Query Match             15.5%;  Score 292.5;  DB 2;  Length 651;
  Best Local Similarity   29.9%;  Pred. No. 1.2e-12;
  Matches   73;  Conservative   48;  Mismatches  102;  Indels   21;  Gaps    7;

Qy         85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLP-------PHKHVARPTEVLAGTQLLYAFF 137
              | |  || :   |:   :|:|     |  :            |:|| |    |:   : ::
Db         30 ASHLITGEKVAIKIMD-KESLGDDLPRVKTEIDAMKNLSHQHVCRLYHVIETPKKIFMVL 88

Qy        138 TR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERK 196
                 |::   : :: |: | || | |||: :|:|: |  |   ||||    :  : :
Db         89 EYCPGGELFDYIIAKDRLTEEEARVFFRQIVSAVAYIHSQGYAHRDLKPENLLIDEDQNL 148

Qy        197 KLVLENLEDSCVLTGPDDSLWDKH-----ACPAYVGPEILSSRASYSGKAADVWSLGVAL 251
              ||:     |  :   |   | | |        |||  ||::  :| | |  ||:||:|| :
Db        149 KLI-----DFGLCAKPKGGL-DYHLMTCCGSPAYAAPELIQGKA-YIGSEADIWSMGVLM 201

Qy        252 FTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILL 311
              : :: |: || |   ::|: || || | :|  || :  |:  ::: :| :|:|  :|
Db        202 YALMCGYLPFDDDNVMVLYKKIMRGKYEIPKWLSPGSVLLLSQMMQVDPKKRITVKHLLN 261

Qy        312 HPWL 315
              ||||
Db        262 HPWL 265


RESULT 4
S72513
FOG2 protein - yeast (Kluyveromyces marxianus var. lactis)
C;Species: Kluyveromyces marxianus var. lactis, Candida sphaerica
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 07-May-1999
C;Accession: S72513
R;Goffrini, P.; Ficarelli, A.; Donnini, C.; Lodi, T.; Puglisi, P.P.; Ferrero, I.
Curr. Genet. 29, 316-326, 1996
A;Title: FOG1 and FOG2 genes, required for the transcriptional activation of glucose-r
A;Reference number: S72513; MUID:96171514; PMID:8598052
A;Accession: S72513
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-602 <GOF>
A;Note: the source is designated as Kluyveromyces lactis
C;Genetics:
A;Gene: FOG2
C;Function:
A;Description: probably involved in the regulation of glucose-repressible gene express
C;Superfamily: AMP-activated protein kinase; protein kinase homology
F;33-286/Domain: protein kinase homology <KIN>

  Query Match             15.4%;  Score 291;  DB 2;  Length 602;
  Best Local Similarity   28.5%;  Pred. No. 1.4e-12;
  Matches   88;  Conservative   51;  Mismatches  138;  Indels   32;  Gaps   10;

Qy         63 ATASRLGPYVLLEP-EEG--GRAYRALHCPTGTEYTCKVYPVQEALAVLE---------P 110
              |    :| | :::   ||  |:   | |  || :   |:   :: ||  :
Db         27 AQGQHIGKYQIIKTLGEGSFGKVKLAYHISTGQKVALKIIN-KKVLAKSDMQGRIERSIS 85

Qy        111 YARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATAL 170
              | ||  | |: :  :|:      :          ::   :  | ::|| ||    |:|: :|:
Db         86 YLRLLRHPHIIKLYDVIKSKDEIIMVIEYAGNELFDYIVQRDKMPEQEARRFFQQIISAV 145

Qy        171 AHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHAC--PAYVGP 228
               :||:| :| ||||    :  :    |:    |  | ::|   |  : | :|  | |  |
Db        146 DYCHRHKIVHRDLKPENLLLDEHLNVKIADFGL--SNIMT---DGNFLKTSCGSPNYAAP 200

Qy        229 EILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPA 288
              |::| :  |:|    |||| || |: ||     || |     :||  |  | | :|  || |
Db        201 EVISGKL-YAGPEVDVWSSGVILYVMLCRRLPFDDESIPVLFKNISNGVYTIPNFLSQGA 259

Qy        289 RCLVRCLLRREPAERLTATGILLHPWLRQD-PMPLAPTRSHLWEAAQVVPDGLGLDEARE 347
               |:: :|   |  |:|   |:   | : | | | | |   | :           |::
Db        260 ASLIKKMLIVNPVNRITVHEIMQDEWFKVDLPDYLVPAESTHQENS----------ESKT 309

Qy        348 EEGDREVVL 356
              |:|   | |
Db        310 EDGGPSVPL 318


RESULT 5
```

S37928
probable purine nucleotide-binding protein YKL101w - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YKL453
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999
C;Accession: S37928; S39084
R;Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzo
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37920
A;Accession: S37928
A;Molecule type: DNA
A;Residues: 1-1518 <CHR>
A;Cross-references: EMBL:Z28101; NID:g486168; PIDN:CAA81941.1; PID:g486169; MIPS:YKL101w
A;Experimental source: strain S288C
R;Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhar
Yeast 9, 1149-1155, 1993
A;Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically loc
protein kinases.
A;Reference number: S39084; MUID:94078677; PMID:8256524
A;Accession: S39084
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-1518 <PAL>
A;Cross-references: EMBL:X71133; NID:g431205; PIDN:CAA50456.1; PID:g431215
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:HSL1
A;Cross-references: SGD:S0001584; MIPS:YKL101w
A;Map position: 11L
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP; P-loop; purine nucleotide binding; serine/threonine-specific protein ki
F;79-369/Domain: protein kinase homology <KIN>
F;79-86/Region: nucleotide-binding motif A (P-loop)
F;87-95/Region: protein kinase ATP-binding motif
F;85/Binding site: ATP/GTP (Lys) #status predicted

Query Match 15.2%; Score 287; DB 2; Length 1518;
Best Local Similarity 26.7%; Pred. No. 7.3e-12;
Matches 86; Conservative 50; Mismatches 128; Indels 58; Gaps 9;

```
Qy          56 PDRATAVATAS-------RLGPYVL---LEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
               ||   :||| |        :||: |    |      ||   | :   ||       |: | ::|
Db          59 PDSTVSVATKSSKRKSRDTVGPWKLGKTLGKGSSGRVRLAKNMETGQLAAIKIVPKKKAF 118

Qy         106 ---------------------------------AVLEPYA--------RLPPHKHVAR 122
                                               :   ||          :|  | :|
Db         119 VHCSNNGTVPNSYSSSMVTSNVSSPSIASREHSNHSQTNPYGIEREIVIMKLISHTNVMA 178

Qy         123 PTEVLAGTQLLYAFFTRTH-GDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLR 181
                 ||       ||        |::   : |: ::|| ||   |:|:   :::||   :  |
Db         179 LFEVWENKSELYLVLEYVDGGELFDYLVSKGKLPEREAIHYFKQIVEGVSYCHSFNICHR 238

Qy         182 DLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKA 241
               |||     :  |:: ::: : :      |  |:  |       | |  |||:  |  | |
Db         239 DLK-PENLLLDKKNRRIKIADF-GMAALELPNKLLKTSCGSPHYASPEIVMGR-PYHGGP 295

Qy         242 ADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPA 301
               :|||| |: || :| || || |      |  |:: | | :|: ||: || |:  :|  :|
Db         296 SDVWSCGIVLFALLTGHLPFNDDNIKKLLLKVQSGKYQMPSNLSSEARDLISKILVIDPE 355

Qy         302 ERLTATGILLHPWLRQ-DPMPL 322
               :|:|   || || ::: | :|:
Db         356 KRITTQEILKHPLIKKYDDLPV 377
```

RESULT 6
T10449
probable serine/threonine-specific protein kinase (EC 2.7.1.-) - cucumber
N;Alternate names: SNF1-related protein kinase
C;Species: Cucumis sativus (cucumber)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C;Accession: T10449
R;Gumpel, N.J.
submitted to the EMBL Data Library, December 1996
A;Reference number: Z17020
A;Accession: T10449
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-504 <GUM>
A;Cross-references: EMBL:Y10036
A;Experimental source: cv. Masterpiece; cotyledon
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threon
C;Superfamily: AMP-activated protein kinase; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;6-260/Domain: protein kinase homology <KIN>

Query Match 15.0%; Score 284; DB 2; Length 504;
Best Local Similarity 28.3%; Pred. No. 3.4e-12;
Matches 86; Conservative 48; Mismatches 116; Indels 54; Gaps 10;

```
Qy          80 GRAYRALHCPTGTEYTCKVY------------PVQEALAVLEPYARLPPHKHVARPTEVL 127
               |:   | |  || :   |:              |:  : :|    ||  | |: |  ||:
Db          20 GKVKIAEHALTGHKVAIKILNRRKIKNLDMEEKVRREIKIL----RLFMHPHIIRLYEVI 75

Qy         128 AGTQLLYAFFTRT-HGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLK-- 184
                  :|       |::   :  : |: | ||   |:|: : : :||:: :| ||||
Db          76 ETPSDIYVVMEYVKSGELFDYIVEKGRLQEDEARNFFQQIISGVEYCHRNMVVHRDLKPE 135

Qy         185 ------LCRFVFADRERKKLVLEN--LEDSCVLTGPDDSLWDKHACPAYVGPEILSSRAS 236
                     |    ||      :: :   |: ||                | |  ||::| :
Db         136 NLLLDSKCNVKIADFGLSNIMRDGHFLKTSC------------GSPNYAAPEVISGKL- 181

Qy         237 YSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLL 296
               |:|    |||| || |: :| |  || |      || ||: | | ||: ||: || |:  :|
Db         182 YAGPEVDVWSCGVILYALLCGTLPFDDENIPNLFKKIKGGIYTLPSHLSSGARELIPSML 241

Qy         297 RREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGD--REV 354
                 :| :|:|   |  |||           ::||     | |    : :|:: : |  :||
Db         242 VVDPMKRITIPEIRQHPWF----------QAHLPRYLAVPPPDT-MQQAKKIDEDILQEV 290

Qy         355 VLYG 358
               | |
Db         291 VKMG 294
```

RESULT 7
T52633
serine/threonine-specific protein kinase (EC 2.7.1.-) AKIN11 [validated] - Arabidopsis
N;Alternate names: SNF1 protein kinase omolog AKIN11
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000
C;Accession: T52633
R;Bhalerao, R.P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Mach
Proc. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999
A;Title: Regulatory interaction of PRL1 WD protein with Arabidopsis SNF1-like protein
A;Reference number: Z25116; MUID:99238528; PMID:10220464
A;Accession: T52633
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-512 <BHA>
A;Cross-references: EMBL:X99279; PIDN:CAA67671.1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: AKIN11
C;Function:
A;Description: EC 2.7.1.-; serine/threonine-specific protein kinase AKIN11 [validated,
complements SNF1 mutations in yeast
C;Superfamily: AMP-activated protein kinase; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

Query Match 15.0%; Score 284; DB 2; Length 512;
Best Local Similarity 28.6%; Pred. No. 3.5e-12;
Matches 87; Conservative 48; Mismatches 115; Indels 54; Gaps 10;

```
Qy          80 GRAYRALHCPTGTEYTCKVY------------PVQEALAVLEPYARLPPHKHVARPTEVL 127
```

```
                |:   | |  || :   |:              |:  : :|    ||  | |: |  ||:
Db         32 GKVKIAEHVVTGHKVAIKILNRRKIKNMEMEEKVRREIKIL----RLFMHPHIIRQYEVI 87

Qy        128 AGTQLLYAFFTRT-HGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLK-- 184
              |  :|       |::   :  : |: | ||   |:|: : : :||:: :| ||||
Db         88 ETTSDIYVVMEYVKSGELFDYIVEKGRLQEDEARNFFQQIISGVEYCHRNMVVHRDLKPE 147

Qy        185 ------LCRFVFADRERKKLVLEN--LEDSCVLTGPDDSLWDKHACPAYVGPEILSSRAS 236
                   |    ||     :: :   |: ||              | |  ||::| :
Db        148 NLLLDSRCNIKIADFGLSNVMRDGHFLKTSC------------GSPNYAAPEVISGKL- 193

Qy        237 YSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLL 296
              |:|   |||| || |: :| |  || |     || ||: | | ||: ||: || |:  :|
Db        194 YAGPEVDVWSCGVILYALLCGTLPFDDENIPNLFKKIKGGIYTLPSHLSSEARDLIPRML 253

Qy        297 RREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARE--EEGDREV 354
              :| :|:|   |. | |           ::||     | |    :::|::  || :||
Db        254 IVDPVKRITIPEIRQHRWP----------QTHLPRYLAVSPPDT-VEQAKKINEEIVQEV 302

Qy        355 VLYG 358
              | |
Db        303 VNMG 306
```

RESULT 8
JC1446
serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana
N;Alternate names: protein kinase SNF1 homolog
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jun-1999
C;Accession: JC1446; S58266; S66334
R;LeGuen, L.; Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.
Gene 120, 249-254, 1992
A;Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein
A;Reference number: JC1446; MUID:93013041; PMID:1339373
A;Accession: JC1446
A;Molecule type: DNA
A;Residues: 1-512 <LEG>
A;Cross-references: GB:M93023; NID:g166599; PIDN:AAA32736.1; PID:g166600
R;Thuemmler, F.; Kirchner, M.; Teuber, R.; Dittrich, P.
submitted to the EMBL Data Library, May 1995
A;Description: Differential accumulation of the transcripts of 22 novel protein kinase g
A;Reference number: S58256
A;Accession: S58266
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 144-198 <THU>
A;Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910
R;Thuemmler, F.; Kirchner, M.; Teuber, R.; Dittrich, P.
Plant Mol. Biol. 29, 551-565, 1995
A;Title: Differential accumulation of the transcripts of 22 novel protein kinase genes i
A;Reference number: S66314; MUID:96123233; PMID:8534852
A;Accession: S66334
A;Molecule type: DNA
A;Residues: 144-198 <TH2>
A;Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910
C;Comment: This enzyme plays an important role in a signal transduction cascade regulati
C;Genetics:
A;Gene: AKin10; AK21
A;Introns: 64/1; 125/3; 186/3; 230/3; 292/3; 322/3; 350/3; 396/3; 475/3
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C;Superfamily: AMP-activated protein kinase; protein kinase homology
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F;17-271/Domain: protein kinase homology <KIN>
F;25-33/Region: protein kinase ATP-binding motif
F;48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F;147,151/Binding site: magnesium (Asn, Asp) #status predicted

  Query Match   14.9%; Score 281; DB 1; Length 512;
  Best Local Similarity   28.7%; Pred. No. 5.6e-12;
  Matches   87; Conservative   46; Mismatches 118; Indels   52; Gaps   9;

```
Qy         80 GRAYRALHCPTGTEYTCKVY-----------PVQEALAVLEPYARLPPHKHVARPTEVL 127
              ||   | |  || :   |:              |:  : :|    ||  | |: |  ||:
Db         31 GRVKIAEHALTGHKVAIKILNRRKIKNMEMEEKVRREIKIL----RLFMHPHIIRLYEVI 86

Qy        128 AGTQLLYAFFTRTH-GDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLK-- 184
                 :|      : |::   :  : |: | ||   |:|: : : :||:: :| ||||
Db         87 ETPTDIYLVMEYVNSGELFDYIVEKGRLQEDEARNFFQQIISGVEYCHRNMVVHRDLKPE 146

Qy        185 ------LCRFVFADRERKKLVLEN--LEDSCVLTGPDDSLWDKHACPAYVGPEILSSRAS 236
                   |    ||     :: :   |: ||              | |  ||::| :
Db        147 NLLLDSKCNVKIADFGLSNIMRDGHFLKTSC------------GSPNYAAPEVISGKL- 192

Qy        237 YSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLL 296
              |:|   |||| || |: :| |  || |     || ||: | | ||: ||  || |:  :|
Db        193 YAGPEVDVWSCGVILYALLCGTLPFDDENIPNLFKKIKGGIYTLPSHLSPGARDLIPRML 252

Qy        297 RREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQV-VPDGLGLDEAREEEGDREVV 355
                :| :|:|   |  |||           ::||     |  || :   :  :||  :||:
Db        253 VVDPMKRVTIPEIRQHPWF----------QAHLPRYLAVPPPDTVQQAKKIDEEILQEVI 302

Qy        356 LYG 358
                |
Db        303 NMG 305
```

RESULT 9
S56719
serine/threonine-specific protein kinase SPK-1 (EC 2.7.1.-) - soybean
C;Species: Glycine max (soybean)
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 18-Jun-1999
C;Accession: S56719
R;Shin, P.G.; Yoon, H.W.; Jeong, Y.H.; Bahk, J.D.; Hong, J.C.; Cho, M.J.
submitted to the EMBL Data Library, January 1993
A;Description: Cloning of a novel protein serine/threonine kinase cDNA from soybean.
A;Reference number: S56719
A;Accession: S56719
A;Molecule type: mRNA
A;Residues: 1-339 <SHI>
A;Cross-references: EMBL:L01453; NID:g169990; PIDN:AAA33979.1; PID:g169991
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;3-261/Domain: protein kinase homology <KIN>
F;11-19/Region: protein kinase ATP-binding motif

  Query Match   14.8%; Score 280; DB 2; Length 339;
  Best Local Similarity   29.3%; Pred. No. 4.2e-12;
  Matches   86; Conservative   35; Mismatches   85; Indels   88; Gaps   11;

```
Qy        117 HKHVARPTEVLAGTQLLYAFFTRTH----------GDMHSLVRSRHRIPEPEAAVLFRQM 166
              | :: |  ||        | | ||          |::   : :  |: | ||   |:|:
Db         58 HPNIIRFKEV---------FLTPTHLAIVLEYAAGGELFERICNAGRLSEDEARFFFQQL 108

Qy        167 ATALAHCHQHGLVLRDLKL--------------CRFVFADRERKKLVLENLEDSCVLTG 211
                : :::||   :  |||||               | | |:    |  :| :   | | |
Db        109 ISGVSYCHSMQICHRDLKLENTLLDGNPAPRLKICDFGFS----KSALLHSQPKSTVGT- 163

Qy        212 PDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF- 270
                      |||: ||:| ||  | || ||||| || |: || | |||:| |    |
Db        164 -----------PAYIAPEVL-SRKEYDGKVADVWSCGVTLYVMLVGAYPFEDPEDPKNFR 211

Qy        271 ---GKIRRGAYALP--AGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD-PMPL-- 322
                 |:|    ||:|    :|   | |: |:    ||:|:: : |  | | | |:: |  :
Db        212 KSIGRIMSVQYAIPDYVRVSKECRHLISCIFVANPAKRISISEIKQHLWFRKNLPREIIE 271

Qy        323 -------------------------APTRSHLWEAAQV-VPDGLGLDEAREE 348
                                      | |: |  | |     | :  ||| ||
Db        272 PERRGYVDHKRTQPSQSVEETMRIIQEARTKIHTGEQAGTGTSDAVHGDEANEE 325
```

RESULT 10
T02306
probable protein kinase [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F13P17.2

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C;Accession: T02306; D84753
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, July 1998
A;Description: Arabidopsis thaliana chromosome II BAC F13P17 genomic sequence.
A;Reference number: Z14657
A;Accession: T02306
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-502 <ROU>
A;Cross-references: EMBL:AC004481; NID:g3337347; PIDN:AAC27394.1; PID:g3337349
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-502 <STO>
A;Cross-references: GB:AE002093; NID:g3337349; PIDN:AAC27394.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g34180; F13P17.2
A;Map position: 2
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
F;55-311/Domain: protein kinase homology <KIN>

Query Match 14.7%; Score 277.5; DB 2; Length 502;
Best Local Similarity 26.0%; Pred. No. 9.6e-12;
Matches 89; Conservative 59; Mismatches 193; Indels 61; Gaps 12;

```
Qy          3 ATPLAAP--------AGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPT 54
              :|||| |         || |:|   :  : | :|  |    |:     || |
Db          6 STPLAIPGPTPIQFMAGLLAR--IVTKNTNKETSTPESPRS-----PRTP---------- 48

Qy         55 APDRATAVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYP------------VQ 102
                : : :     :|   ||      : | | :  :| :    ||              ::
Db         49 ---QGSILMDKYEIGK--LLGHGSFAKVYLARNIHSGEDVAIKVIDKEKIVKSGLAGHIK 103

Qy        103 EALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTH-GDMHSLVRSRHRIPEPEAAV 161
              :::|    |   | ::    ||:|    :|        |:::: | :| |: |  |
Db        104 REISIL----RRVRHPYIVHLLEVMATKTKIYIVMEYVRGGELYNTV-ARGRLREGTARR 158

Qy        162 LFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHA 221
              |:|: :::| ||  |:  |||||    :  |:   |:          |:   : |  :
Db        159 YFQQLISSVAFCHSRGVYHRDLKLENLLLDDKGNVKV------SDFGLSVVSEQLKQEGI 212

Qy        222 C------PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRR 275
              |      |||: ||:| :|  | |  ||:|| || || ::||: || |   :::: || :
Db        213 CQTFCGTPAYLAPEVL-TRKGYEGAKADIWSCGVILFVLMAGYLPFDDKNILVMYTKIYK 271

Qy        276 GAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQ 317
              | :  |   |     ||  :|   |  |:|   |: | | ::
Db        272 GQFKCPKWFSPELARLVTRMLDTNPDTRITIPEIMKHRWFKK 313
```

RESULT 11
E85362
hypothetical protein AT4g30960 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C;Accession: E85362
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: E85362
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <STO>
A;Cross-references: GB:NC_001268; NID:g7269998; PIDN:CAB79814.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g30960
A;Map position: 4
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homo

Query Match 14.6%; Score 277; DB 2; Length 441;
Best Local Similarity 29.4%; Pred. No. 9e-12;
Matches 77; Conservative 39; Mismatches 124; Indels 22; Gaps 4;

```
Qy         81 RAYRALHCPTGTEYTCKVY-----------PVQEALAVLEPYARLPPHKHVARPTEVLA 128
              : | | :  ||      ||             ::  ::|:    |:  | ::    ||:|
Db         37 KVYHARNIQTGKSVAMKVVGKEKVVKVGMVDQIKREISVM----RMVKHPNIVELHEVMA 92

Qy        129 GTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRF 188
                :|       |       :: |: |  | | |:|: :|:  ||  |:  ||||
Db         93 SKSKIYFAMELVRGGELFAKVAKGRLREDVARVYFQQLISAVDFCHSRGVYHRDLKPENL 152

Qy        189 VFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLG 248
              :  :    |:    |          | |      |||| ||::  :  | |  ||:|| |
Db        153 LLDEEGNLKVTDFGLSAFTEHLKQDGLLHTTCGTPAYVAPEVILKK-GYDGAKADLWSCG 211

Qy        249 VALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATG 308
              | || :|||: ||||   | :: || || :  |  ||: || ||  ||   |  |:|
Db        212 VILFVLLAGYLPFQDDNLVNMYRKIYRGDFKCPGWLSSDARRLVTKLLDPNPNTRITIEK 271

Qy        309 ILLHPWLRQDPM-----PLAPT 325
              ::  || ::        |:| |
Db        272 VMDSPWFKKQATRSRNEPVAAT 293
```

RESULT 12
A56009
serine/threonine-specific protein kinase (EC 2.7.1.-) NPK5 - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C;Accession: A56009
R;Muranaka, T.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994
A;Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces c
tase of Saccharomyces cerevisiae.
A;Reference number: A56009; MUID:94217693; PMID:8164654
A;Accession: A56009
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-511 <MUR>
A;Cross-references: GB:D26602; NID:g496384; PIDN:BAA05649.1; PID:g496385
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threon
C;Superfamily: AMP-activated protein kinase; protein kinase homology
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kina
F;17-271/Domain: protein kinase homology <KIN>
F;25-33/Region: protein kinase ATP-binding motif
F;48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F;147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 14.6%; Score 277; DB 1; Length 511;
Best Local Similarity 28.3%; Pred. No. 1.1e-11;
Matches 86; Conservative 49; Mismatches 115; Indels 54; Gaps 10;

```
Qy         80 GRAYRALHCPTGTEYTCKVY-----------PVQEALAVLEPYARLPPHKHVARPTEVL 127
              |:   | |  || :  |:               |:  : :|    ||  | |: |  ||:
Db         31 GKVKIAEHTLTGHKVAVKILNRRKIKNMEMEEKVRREIKIL----RLFMHPHIIRLYEVV 86

Qy        128 AGTQLLYAFFTRT-HGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLC 186
                :|       |::  :  : |: | ||    |:|: : : :||:: :| ||||
Db         87 ETPSDIYVVMEYVKSGELFDYIVEKGRLQEDEARKFFQQIISGVEYCHRNMVVHRDLKPE 146

Qy        187 RFVFADRERKKLV---LEN-------LEDSCVLTGPDDSLWDKHACPAYVGPEILSSRAS 236
              :   :   |:    | |        |: ||             | |  ||::| :
Db        147 NLLLDSKWNVKIADFGLSNIMRDGHFLKTSC------------GSPNYAAPEVISGKL- 192

Qy        237 YSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLL 296
```

```
Db        193 YAGPEVDVWSCGVILYALLCGTLPFDDENIPNLFKKIKGGMISLPSHLSAGARDLIPRML 252

Qy        297 RREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARBEEGD--REV 354

Db        253 IVDPMKRMTIPEIRMHPWF----------QAHLPRYLAVPPPDT-MQQAKKIDEDILQEV 301

Qy        355 VLYG 358

Db        302 VKRG 305


RESULT 13
A26030
serine/threonine-specific protein kinase (EC 2.7.1.-) SNF1 - yeast (Saccharomyces cerevi
N;Alternate names: protein YDR477w
C;Species: Saccharomyces cerevisiae
C;Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #text_change 23-Mar-2001
C;Accession: A26030; S69644
R;Celenza, J.L.; Carlson, M.
Science 233, 1175-1180, 1986
A;Title: A yeast gene that is essential for release from glucose repression encodes a pr
A;Reference number: A26030; MUID:86289463; PMID:3526554
A;Accession: A26030
A;Molecule type: DNA
A;Residues: 1-633 <CEL>
A;Cross-references: EMBL:M13971; NID:g172629; PIDN:AAA35058.1; PID:g172630
R;Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A;Description: The sequence of S. cerevisiae cosmids 9410, 8035, 8166, and 9787.
A;Reference number: S69554
A;Accession: S69644
A;Molecule type: DNA
A;Residues: 1-633 <DIE>
A;Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64904.1; PID:g927732; GSPDB:GN000
C;Genetics:
A;Gene: SGD:SNF1; MIPS:YDR477w
A;Cross-references: SGD:S0002885; MIPS:YDR477w
A;Map position: 4R
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A;Note: required for expression of glucose-repressed genes in response to glucose depriv
C;Superfamily: AMP-activated protein kinase; protein kinase homology
C;Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransfe
F;53-306/Domain: protein kinase homology <KIN>
F;61-69/Region: protein kinase ATP-binding motif
F;84,103,177,179/Active site: Lys, Glu, Asp, Lys #status predicted
F;182,186/Binding site: magnesium (Asn, Asp) #status predicted
F;210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predicte

  Query Match             14.5%;  Score 274.5;  DB 1;  Length 633;
  Best Local Similarity   27.8%;  Pred. No. 2e-11;
  Matches   88;  Conservative   53;  Mismatches  138;  Indels   37;  Gaps   11;

Qy         51 SPPTAPDRATAVATASRLGPYVLLEP-EEG--GRAYRALHCPTGTEYTCKVYPVQEALAV 107

Db         35 SNSTLNNPKSSLADGAHIGNYQIVKTLGEGSFGKVKLAYHTTTGQKVALKIIN-KKVLAK 93

Qy        108 LE---------PYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPE 158

Db         94 SDMQGRIEREISYLRLLRHPHIIKLYDVIKSKDEIIMVIEYAGNELFDYIVQRDKMSEQE 153

Qy        159 AAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD 218

Db        154 ARRFFQQIISAVEYCHRHKIVHRDLKPENLLLDEHLNVKIADFGL--SNIMT---DGNFL 208

Qy        219 KHAC--PAYVGPEILSSRASYSGKAADVWSLGVALPTMLAGHYPFQDSEPVLLFGKIRRG 276

Db        209 KTSCGSPNYAAPEVISGKL-YAGPEVDVWSCGVILYVMLCRRLPFDDESIPVLFKNISNG 267

Qy        277 AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD-PMPLAPTRSHLWEAAQV 335

Db        268 VYTLPKFLSPGAAGLIKRMLIVNPLNRISIHEIMQDDWFKVDLPEYLLP---------- 316

Qy        336 VPDGLGLDEAREEEGD 351

Db        317 -PD---LKPHPEEENE 328


RESULT 14
S60304
serine/threonine-specific protein kinase (EC 2.7.1.-) BKIN12 (version 2) - barley
C;Species: Hordeum vulgare (barley)
C;Date: 19-Mar-1997 #sequence_revision 15-Aug-1997 #text_change 11-Jun-1999
C;Accession: S60304; S24579
R;Halford, N.G.; Vicente-Carbajosa, J.; Sabelli, P.A.; Shewry, P.R.; Hannappel, U.; Kr
Plant J. 2, 791-797, 1992
A;Title: Molecular analyses of a barley multigene family homologous to the yeast prote
A;Reference number: S60303; MUID:93258420; PMID:1302632
A;Accession: S60304
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-513 <HAL>
A;Cross-references: EMBL:X65604
R;Halford, N.G.
submitted to the EMBL Data Library, April 1992
A;Reference number: S24578
A;Accession: S24579
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-61,'A',63-513 <HA2>
A;Cross-references: EMBL:X65604; NID:g18933; PIDN:CAA46554.1; PID:g18934
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threon
C;Superfamily: AMP-activated protein kinase; protein kinase homology
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kina
F;15-272/Domain: protein kinase homology <KIN>
F;23-31/Region: protein kinase ATP-binding motif
F;46,65,143,145/Active site: Lys, Glu, Asp, Lys #status predicted
F;148,152/Binding site: magnesium (Asn, Asp) #status predicted

  Query Match             14.5%;  Score 274;  DB 1;  Length 513;
  Best Local Similarity   31.7%;  Pred. No. 1.7e-11;
  Matches   79;  Conservative   42;  Mismatches  102;  Indels   26;  Gaps    7;

Qy        117 HKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQ 175

Db         77 HPHIIRVYEVIETPKDIFVVMEYCNNGELLDYIIENGRLQEDEARRIFQQILAGVEYCHR 136

Qy        176 HGLVLRDLKLCRFVFADRERKKLV---LENLEDSCVLTGPDDSLWDKHACPA--YVGPEI 230

Db        137 IMVVHRDLKPENLLLDSRYNVKLADFGLSNV--------MRDGHFLKTSCGSLNYAAPEI 188

Qy        231 LSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARC 290

Db        189 ISSKL-YAGPEVDVWSCGVVLYALLCGSVPFDDDNIPSLFRKIKGGTYILPSYLSDSARD 247

Qy        291 LVRCLLRREPAERLTATGILLHPWLRQD-PMPLAPTRSHLWEAAQVVPDGLGLDEAREEE 349

Db        248 LIPKLLNIDPMKRITFHEIRVHPWFKNHLPCYLAVPPPYKEQQAKMI----------DED 297

Qy        350 GDREVVLYG 358

Db        298 ILREVVNLG 306


RESULT 15
A34366
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) II delta chain - rat
N;Contains: Ca2+/calmodulin-dependent protein kinase II delta chain, various splice fo
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Jun-1990 #sequence_revision 08-Jun-1990 #text_change 11-Jun-1999
C;Accession: A34366; S39023; S39024; S39025; S39026; S39027; C47170
R;Tobimatsu, T.; Fujisawa, H.
J. Biol. Chem. 264, 17907-17912, 1989
A;Title: Tissue-specific expression of four types of rat calmodulin-dependent protein
```

A;Reference number: A34366; MUID:90036861; PMID:2553697
A;Accession: A34366
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-533 <TOB>
A;Cross-references: GB:J05072; NID:g203266; PIDN:AAA40866.1; PID:g203267
R;Mayer, P.; Moehlig, M.; Schatz, H.; Pfeiffer, A.
FEBS Lett. 333, 315-318, 1993
A;Title: New isoforms of multifunctional calcium/calmodulin-dependent protein kinase II.
A;Reference number: S39023; MUID:94039784; PMID:8224201
A;Accession: S39023
A;Molecule type: mRNA
A;Residues: 318-328,363-371,504-533 <MAY>
A;Accession: S39024
A;Molecule type: mRNA
A;Residues: 318-371,504-511,'N' <MA2>
A;Accession: S39025
A;Molecule type: mRNA
A;Residues: 318-328,363-371,504-511,'N' <MA3>
A;Accession: S39026
A;Molecule type: mRNA
A;Residues: 318-349,364-371,504-533 <MA4>
A;Accession: S39027
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 318-349,364-371,504-511,'N' <MA5>
R;Schworer, C.M.; Rothblum, L.I.; Thekkumkara, T.J.; Singer, H.A.
J. Biol. Chem. 268, 14443-14449, 1993
A;Title: Identification of novel isoforms of the delta subunit of Ca2+/calmodulin-depend
A;Reference number: A47170; MUID:93300844; PMID:8390994
A;Accession: C47170
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 314-349,364-368 <SCH>
A;Cross-references: GB:L13408; NID:g349086
A;Experimental source: skeletal muscle
A;Note: sequence extracted from NCBI backbone (NCBIN:134450, NCBIP:134453)
C;Complex: heteromultimer composed of 10-12 alpha, beta, gamma, and delta chains general
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A;Note: acts on a variety of intracellular proteins; gamma and delta chains are expresse
C;Superfamily: Ca2+/calmodulin-dependent protein kinase II; protein kinase homology
C;Keywords: ATP; autophosphorylation; calmodulin binding; phosphoprotein; phosphotransfe
F;12-272/Domain: protein kinase homology <KIN>
F;20-28/Region: protein kinase ATP-binding motif
F;287-311/Region: calmodulin binding #status predicted
F;43,61,136,138/Active site: Lys, Glu, Asp, Lys #status predicted
F;287/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status predicte

Query Match 14.5%; Score 273.5; DB 1; Length 533;
Best Local Similarity 28.8%; Pred. No. 1.9e-11;
Matches 79; Conservative 53; Mismatches 123; Indels 19; Gaps 7;

```
Qy         61 AVATASRLGPYVLLEPEEGGRAY----RALHCPTGTEYTCKVYPVQEALA----VLEPYA 112
              :  | :|       |  | |  |:    | :  ||| ||  |:   ::  |    ||  |
Db          3 STTTCTRFTDEYQLFEELGKGAFSVVRRCMKIPTGQEYAAKIINTKKLSARDHQKLEREA 62

Qy        113 ---RLPPHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMAT 168
                 ||  | :: |  : ::     |  |   | |::   : :|    | :|:   :|:
Db         63 RICRLLKHPNIVRLHDSISEEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIQQILE 122

Qy        169 ALAHCHQHGLVLRDLKLCRFVFADRERKKLV-LENLEDSCVLTGPDDSLWDKHACPAYVG 227
              :: ||| :|:| ||||    : | : :   | | :   :  : |   : :     | |:
Db        123 SVNHCHLNGIVHRDLKPENLLLASKSKGAAVKLADFGLAIEVQGDQQAWFGFAGTPGYLS 182

Qy        228 PEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPA----G 283
              ||:|  |    ||  |:|: || |: :| |: || | :   |: :|: |||  |:
Db        183 PEVL--RKDPYGKPVDMWACGVILYILLVGYPPFWDEDQHRLYQQIKAGAYDFPSPEWDT 240

Qy        284 LSAPARCLVRCLLRREPAERLTATGILLHPWLRQ 317
              ::  |: |:  :|   ||:|:||:  | |||: |
Db        241 VTPEAKDLINKMLTINPAKRITASEALKHPWICQ 274
```

Search completed: January 15, 2004, 14:56:52
Job time : 21 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:57:57 ; Search time 38 Seconds
(without alignments)
1926.428 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 1891
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1887	99.8	358	9	US-09-799-875-8	Sequence 8, Appli
2	1675	88.6	360	15	US-10-024-828-9	Sequence 9, Appli
3	1102	58.3	233	9	US-09-925-301-1102	Sequence 1102, Ap
4	681.5	36.0	269	12	US-10-291-172-304	Sequence 304, App
5	659.5	34.9	290	12	US-10-291-172-680	Sequence 680, App
6	578.5	30.6	206	15	US-10-228-263-2	Sequence 2, Appli
7	408	21.6	138	9	US-09-864-761-45767	Sequence 45767, A
8	354	18.7	153	9	US-09-925-301-1367	Sequence 1367, Ap
9	304.5	16.1	373	15	US-10-026-021-4	Sequence 4, Appli
10	297.5	15.7	778	12	US-10-354-358-92	Sequence 92, Appl
11	297.5	15.7	778	12	US-10-116-326-2	Sequence 2, Appli
12	290.5	15.4	607	12	US-09-769-970-15	Sequence 15, Appl
13	290.5	15.4	607	12	US-10-204-041-16	Sequence 16, Appl
14	290.5	15.4	607	15	US-10-108-580-2	Sequence 2, Appli
15	287	15.2	1518	10	US-09-801-368-152	Sequence 152, App
16	287	15.2	1518	12	US-10-369-493-22243	Sequence 22243, A
17	283.5	15.0	651	9	US-09-870-937-10	Sequence 10, Appl
18	283.5	15.0	651	10	US-09-974-298-112	Sequence 112, App
19	283.5	15.0	651	12	US-10-354-358-8	Sequence 8, Appli
20	283.5	15.0	651	12	US-10-295-027-506	Sequence 506, App
21	283.5	15.0	651	12	US-10-173-999-127	Sequence 127, App
22	280	14.8	420	12	US-10-369-493-3864	Sequence 3864, Ap
23	278	14.7	290	9	US-09-867-550-1772	Sequence 1772, Ap
24	274.5	14.5	446	9	US-09-824-735-2	Sequence 2, Appli
25	274.5	14.5	633	9	US-09-824-735-3	Sequence 3, Appli
26	274.5	14.5	633	10	US-09-801-368-338	Sequence 338, App
27	274.5	14.5	633	12	US-10-369-493-1696	Sequence 1696, Ap
28	273.5	14.5	668	14	US-10-054-579-2	Sequence 2, Appli
29	273.5	14.5	668	15	US-10-195-072-2	Sequence 2, Appli
30	273.5	14.5	668	15	US-10-195-071-2	Sequence 2, Appli
31	273.5	14.5	674	15	US-10-283-247-2	Sequence 2, Appli
32	271.5	14.4	664	12	US-10-288-798-18	Sequence 18, Appl
33	269.5	14.3	674	10	US-09-842-582-2	Sequence 2, Appli
34	269.5	14.3	674	15	US-10-283-247-7	Sequence 7, Appli
35	269.5	14.3	674	15	US-10-283-247-8	Sequence 8, Appli
36	268	14.2	622	12	US-10-369-493-5806	Sequence 5806, Ap
37	266.5	14.1	499	12	US-10-354-358-88	Sequence 88, Appl
38	266	14.1	516	12	US-09-820-790-2	Sequence 2, Appli
39	266	14.1	542	12	US-09-820-790-4	Sequence 4, Appli
40	265.5	14.0	703	12	US-10-116-326-6	Sequence 6, Appli
41	263	13.9	1142	12	US-10-369-493-1700	Sequence 1700, Ap
42	261.5	13.8	436	14	US-10-006-611-2	Sequence 2, Appli
43	261.5	13.8	891	12	US-10-369-493-2533	Sequence 2533, Ap
44	260.5	13.8	786	11	US-09-823-187-91	Sequence 91, Appl
45	260.5	13.8	786	12	US-10-231-913-118	Sequence 118, App

ALIGNMENTS

RESULT 1
US-09-799-875-8
; Sequence 8, Application US/09799875
; Patent No. US20020034780A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. US20020034780A1el Human Protein Kinases and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-875-8

Query Match 99.8%; Score 1887; DB 9; Length 358;
Best Local Similarity 99.7%; Pred. No. 2.1e-157;
Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db         61 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300

Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358


RESULT 2
US-10-024-828-9
; Sequence 9, Application US/10024828
; Publication No. US20030036051A1
; GENERAL INFORMATION:
;  APPLICANT: Virca, Duke
;  APPLICANT:  Bird, Timothy A.
;  APPLICANT:  Anderson, Dirk M.
;  APPLICANT:  Marken, John S.
;  TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase
;  TITLE OF INVENTION:  Functions
;  FILE REFERENCE: 2877-US
;  CURRENT APPLICATION NUMBER: US/10/024,828
;  CURRENT FILING DATE:  2001-12-18
;  PRIOR APPLICATION NUMBER: US/09/509,902A
;  PRIOR FILING DATE: 1999-08-03
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 360
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-024-828-9

  Query Match            88.6%;  Score 1675;  DB 15;  Length 360;
  Best Local Similarity   99.7%;  Pred. No. 8.8e-139;
  Matches  317;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 102

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
                 ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        103 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 162

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 222

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 282

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 342

Qy        301 AERLTATGILLHPWLRQD 318
                 ||||||||||||||||||
Db        343 AERLTATGILLHPWLRQD 360


RESULT 3
US-09-925-301-1102
; Sequence 1102, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1102
;   LENGTH: 233
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-301-1102

  Query Match            58.3%;  Score 1102;  DB 9;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 1e-88;
  Matches  208;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        151 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 210
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 85

Qy        211 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 270
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 145

Qy        271 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 330
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 205

Qy        331 EAAQVVPDGLGLDEAREEEGDREVVLYG 358
                 ||||||||||||||||||||||||||||
Db        206 EAAQVVPDGLGLDEAREEEGDREVVLYG 233


RESULT 4
US-10-291-172-304
; Sequence 304, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-045
;  CURRENT APPLICATION NUMBER: US/10/291,172
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 09/693,267
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/665,363
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: 09/616,847
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 09/596,193
;  PRIOR FILING DATE: 2000-06-17
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 304
;   LENGTH: 269
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-291-172-304

  Query Match            36.0%;  Score 681.5;  DB 12;  Length 269;
  Best Local Similarity   51.8%;  Pred. No. 1.1e-51;
  Matches  133;  Conservative   38;  Mismatches   79;  Indels    7;  Gaps    2;

Qy         88 CP------TGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTH 141
```

```
Db          2 CPGRCASTLGRRVRCKVPPIKHYQDKIRPYIQLPSHSNITGIVEVILGETKAYVFPEKDF 61

Qy        142 GDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLE 201
              ||||| |||| |: | ||| ||:|: :|:|||||  :|| |||| :|||:  || :| ||
Db         62 GDMHSYVRSRKRLREEEAARLFKQIVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLE 121

Qy        202 NLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPF 261
              :|||: :: | ||:| ||| ||||| ||||::  :|||||||||||| |:|:| | |||
Db        122 SLEDTHIMKGEDDALSDKHGCPAYVSPEILNTTGTYSGKAADVWSLGVMLYTLLVGRYPF 181

Qy        262 QDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMP 321
              ||:|  || ||||| : :|  :|  ||||:| ||||||:|||||  ||||||      |
Db        182 HDSDPSALFSKIRRGQFCIPEHISPKARCLIRSLLRREPSERLTAPEILLHPWFESVLEP 241

Qy        322 LAPTRSHLWEAAQVVPD 338
                    | :  : |:||:
Db        242 -GYIDSEIGTSDQIVPE 257


RESULT 5
US-10-291-172-680
; Sequence 680, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-045
;  CURRENT APPLICATION NUMBER: US/10/291,172
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 09/693,267
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/665,363
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: 09/616,847
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 09/596,193
;  PRIOR FILING DATE: 2000-06-17
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 680
;   LENGTH: 290
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-291-172-680

  Query Match           34.9%;  Score 659.5;  DB 12;  Length 290;
  Best Local Similarity  51.2%;  Pred. No. 1e-49;
  Matches  127;  Conservative   42;  Mismatches   78;  Indels    1;  Gaps    1;

Qy         91 GTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDMHSLVRS 150
              |  :: :|:|::     : || :|| | ::    ||: |   | || :  |||||  |||
Db         22 GPPFSLQVFPIKHYQDKIRPYIQLPSHSNITGIVEVILGETKAYVFPEKDFGDMHSYVRS 81

Qy        151 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 210
              | |: | ||| ||:|: :|:|||||  :|| |||| :|||:  || :| ||:|||: ::
Db         82 RKRLREEEAARLFKQIVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLESLEDTHIMK 141

Qy        211 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 270
              | ||:| ||| ||||| ||||::  :|||||||||: || |:|:| | ||| ||:|  ||
Db        142 GEDDALSDKHGCPAYVSPEILNTTGTYSGKAADVWTGGVMLYTLLVGRYPFHDSDPSALF 201

Qy        271 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 330
              |||||  : :|  :|  ||||:| ||||||:|||||  ||||||      |    | :
Db        202 SKIRRGQFCIPEHISPKARCLIRSLLRREPSERLTAPEILLHPWFESVLEP-GYIDSEIG 260

Qy        331 EAAQVVPD 338
                : |:||:
Db        261 TSDQIVPE 268


RESULT 6
US-10-228-263-2
; Sequence 2, Application US/10228263
; Publication No. US20030099985A1
; GENERAL INFORMATION:
;  APPLICANT: LI, Jing
;  TITLE OF INVENTION: AMPLIFIED GENE INVOLVED IN CANCER
;  FILE REFERENCE: 38002-0034
;  CURRENT APPLICATION NUMBER: US/10/228,263
;  CURRENT FILING DATE:  2002-12-04
;  PRIOR APPLICATION NUMBER: US 60/330,797
;  PRIOR FILING DATE: 2001-10-31
;  PRIOR APPLICATION NUMBER: US 60/314,655
;  PRIOR FILING DATE: 2001-08-27
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 206
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-228-263-2

  Query Match           30.6%;  Score 578.5;  DB 15;  Length 206;
  Best Local Similarity  57.9%;  Pred. No. 8.3e-43;
  Matches  113;  Conservative   29;  Mismatches   52;  Indels    1;  Gaps    1;

Qy        144 MHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENL 203
              ||| |||| |: | ||| ||:|: :|:|||||  :|| |||| :|||:  || :| ||:|
Db          1 MHSYVRSRKRLREEEAARLFKQIVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLESL 60

Qy        204 EDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQD 263
              ||: :: | ||:| ||| ||||| ||||::  :|||||||||||| |:|:| | ||| |
Db         61 EDTHIMKGEDDALSDKHGCPAYVSPEILNTTGTYSGKAADVWSLGVMLYTLLVGRYPFHD 120

Qy        264 SEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLA 323
              |:|  || ||||| : :|  :|  ||||:| ||||||:|||||  ||||||      |
Db        121 SDPSALFSKIRRGQFCIPEHISPKARCLIRSLLRREPSERLTAPEILLHPWFESVLEP-G 179

Qy        324 PTRSHLWEAAQVVPD 338
                  | :  : |:||:
Db        180 YIDSEIGTSDQIVPE 194


RESULT 7
US-09-864-761-45767
; Sequence 45767, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45767
;   LENGTH: 138
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC009486.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.93
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
;   OTHER INFORMATION: SWISSPROT HIT: O74536, EVALUE 1.00e-18
;   OTHER INFORMATION: EST_HUMAN HIT: BE897149.1, EVALUE 2.00e-71
US-09-864-761-45767

  Query Match             21.6%;  Score 408;  DB 9;  Length 138;
  Best Local Similarity   63.1%;  Pred. No. 4.6e-28;
  Matches   77;  Conservative   15;  Mismatches   30;  Indels    0;  Gaps    0;

Qy        197 KLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLA 256
              :: ||:|||: :| | |||| ||| ||||| ||||:: ||||||||||||| |:|||
Db          2 RVKLESLEDAYILRGDDDSLSDKHGCPAYVSPEILNTSGSYSGKAADVWSLGVMLYTMLV 61

Qy        257 GHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLR 316
              | ||| | ||  || ||||| : :|  ||  |:||:| :|||||:||||:  || |||
Db         62 GRYPFHDIEPSSLFSKIRRGQFNIPETLSPKAKCLIRSILRREPSERLTSQEILDHPWFS 121

Qy        317 QD 318
              |
Db        122 TD 123


RESULT 8
US-09-925-301-1367
; Sequence 1367, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1367
;   LENGTH: 153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (136)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (138)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (141)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (142)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (143)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (152)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1367

  Query Match             18.7%;  Score 354;  DB 9;  Length 153;
  Best Local Similarity   93.3%;  Pred. No. 2.9e-23;
  Matches   70;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 133

Qy         61 AVATASRLGPYVLLE 75
              || | ||   |||||
Db        134 AVXTXSRXXXYVLLE 148


RESULT 9
US-10-026-021-4
; Sequence 4, Application US/10026021
; Publication No. US20030027756A1
; GENERAL INFORMATION:
;  APPLICANT: Hitoshi, Yasumichi
;  APPLICANT:  Demo, Susan
;  APPLICANT:  Jenkins, Yonchu
;  APPLICANT:  Rigel Pharmaceuticals, Inc.
;  TITLE OF INVENTION: SAK:  Modulation of Cellular Proliferation for
;  TITLE OF INVENTION:  Treatment of Cancer
;  FILE REFERENCE: 021044-001210US
;  CURRENT APPLICATION NUMBER: US/10/026,021
;  CURRENT FILING DATE:  2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/309,632
;  PRIOR FILING DATE: 2001-08-01
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (1)..(373)
;   OTHER INFORMATION: human FNK mitotic kinase kinase domain
US-10-026-021-4

  Query Match             16.1%;  Score 304.5;  DB 15;  Length 373;
  Best Local Similarity   28.9%;  Pred. No. 2.1e-18;
  Matches  101;  Conservative   48;  Mismatches  134;  Indels   67;  Gaps   15;

Qy         29 RPVQKRAR-----SGPQPRLPPCL----------LPLSPP----TAPDRATAVATASRL 68
              || |: |      :|| |  ||             || | |    | ||          |
Db         11 RPFQRTAAATAPPAGPGP--PPSALRGPELEMLAGLPTSDPGRLITDPRSGRTYLKGRLL 68
```

```
Qy          69 GPYVLLEPEEGG--RAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLP---------- 115
               |         :||  | | |     ||: |  || |       |::|: |
Db          69 G--------KGGPARCYEATDTETGSAYAVKVIPQSR---VVKPHQREKILNEIELHRDL 117

Qy         116 PHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCH 174
                |:|: | :          :| |     :   :  : ::|| : |||     ||: : | : |
Db         118 QHRHIVRFSHHFEDADNIYIFLELCSRKSLAHIWKARHTLLEPEVRYYLRQILSGLKYLH 177

Qy         175 QHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHAC--PAYVGPEILS 232
               | |:: |||||  |     |  :| : :   :  |  |:      |  |  | || ||:|
Db         178 QRGILHRDLKLGNFFIT--ENMELKVGDFGLAARLEPPEQR--KKTICGTPNYVAPEVLL 233

Qy         233 SRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLV 292
                :    | |||||||  ::|:| |  ||: ::    :  |::  | ||| || ||| |:
Db         234 RQG--HGPEADVWSLGCVMYTLLCGSPPFETADLKETYRCIKQVHYTLPASLSLPARQLL 291

Qy         293 RCLLRREPAERLTATGILLHPWLRQ----DPMPLAPTRSHLWEAAQVVPD 338
                 :||  | :| :   || | :  :    | :|:: :   |||
Db         292 AAILRASPRDRPSIDQILRHDFFTKGYTPDRLPIS--------SCVTVPD 333
```

RESULT 10
US-10-354-358-92
; Sequence 92, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MPI02-020P1RNOMNIM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-354-358-92

Query Match 15.7%; Score 297.5; DB 12; Length 778;
Best Local Similarity 27.4%; Pred. No. 2.2e-17;
Matches 98; Conservative 54; Mismatches 143; Indels 63; Gaps 14;

```
Qy          44 PPCLLPLSPPTAPDRATAVATASRLGPYVL---LEPEEGGRAYRALHCPTGTEYTCKV-- 98
               |   ||   |  |  |  |       ||| |   |   : |     :|| || :   |:
Db          13 PAYHLPHPHPHPPQHAQYV------GPYRLEKTLGKGQTGLVKLGVHCITGQKVAIKIVN 66

Qy          99 ---------YPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLV 148
                        |:  :|:|    :|  | || :  :|    : ||      : |::   :
Db          67 REKLSESVLMKVEREIAIL----KLIEHPHVLKLHDVYENKKYLYLVLEHVSGGELFDYL 122

Qy         149 RSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCV 208
                 : |:   ||   |||: :||  || : :  ||||     :    |:  : : :   : :
Db         123 VKKGRLTPKEARKFFRQIVSALDFCHSYSICHRDLKPENLLL--DEKNNIRIADFGMASL 180

Qy         209 LTGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEP 266
                 |  ||| :  :|  | |  ||::     | |: ||:|| || || :| |  || |
Db         181 QVG--DSLLET-SCGSPHYACPEVIKGE-KYDGRRADMWSCGVILFALLVGALPFDDDNL 236

Qy         267 VLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWL---RQDPMPL- 322
                |  |::|| : :|  :    : |:| ::  || :||:   |  |||    : :| |
Db         237 RQLLEKVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKRLSLEQIQKHPWYLGGKHEPDPCL 296

Qy         323 --APTRSHLWEAAQVVPDGLGLD--------------------EARBEEGDREVVLY 357
                 || |     | : :|   ||                      | | || ::| ::|
Db         297 EPAPGRR---VAMRSLPSNGELDPDVLESMASLGCFRDRERLHRELRSEEENQEKMIY 351
```

RESULT 11
US-10-116-326-2
; Sequence 2, Application US/10116326
; Publication No. US20030166889A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US20030166889A1el Human Kinases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0332-USA
; CURRENT APPLICATION NUMBER: US/10/116,326
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,036
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 778
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-116-326-2

Query Match 15.7%; Score 297.5; DB 12; Length 778;
Best Local Similarity 27.4%; Pred. No. 2.2e-17;
Matches 98; Conservative 54; Mismatches 143; Indels 63; Gaps 14;

```
Qy          44 PPCLLPLSPPTAPDRATAVATASRLGPYVL---LEPEEGGRAYRALHCPTGTEYTCKV-- 98
               |   ||   |  |  |  |       ||| |   |   : |     :|| || :   |:
Db          13 PAYHLPHPHPHPPQHAQYV------GPYRLEKTLGKGQTGLVKLGVHCITGQKVAIKIVN 66

Qy          99 ---------YPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLV 148
                        |:  :|:|    :|  | || :  :|    : ||      : |::   :
Db          67 REKLSESVLMKVEREIAIL----KLIEHPHVLKLHDVYENKKYLYLVLEHVSGGELFDYL 122

Qy         149 RSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCV 208
                 : |:   ||   |||: :||  || : :  ||||     :    |:  : : :   : :
Db         123 VKKGRLTPKEARKFFRQIVSALDFCHSYSICHRDLKPENLLL--DEKNNIRIADFGMASL 180

Qy         209 LTGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEP 266
                 |  ||| :  :|  | |  ||::     | |: ||:|| || || :| |  || |
Db         181 QVG--DSLLET-SCGSPHYACPEVIKGE-KYDGRRADMWSCGVILFALLVGALPFDDDNL 236

Qy         267 VLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWL---RQDPMPL- 322
```

```
Db        237 RQLLEKVKRGVPHMPHFIPPDCQSLLRGMIEVEPEKRLSLEQIQKHPWYLGGKHEPDPCL 296

Qy        323 --APTRSHLWEAAQVVPDGLGLD--------------------EAREEEGDREVVLY 357
Db        297 EPAPGRR---VAMRSLPSNGELDPDVLESMASLGCFRDRERLHRELRSEEENQEKMIY 351


RESULT 12
US-09-769-970-15
; Sequence 15, Application US/09769970
; Publication No. US20030170219A1
;   GENERAL INFORMATION:
;        APPLICANT: Bandman, Olga
;                   Hillman, Jennifer L.
;                   Corley, Neil C.
;                   Guegler, Karl G.
;                   Lal, Preeti
;                   Goli, Surya K.
;                   Shah, Purvi
;        TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
;                            KINASES
;        NUMBER OF SEQUENCES: 21
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/769,970
;             FILING DATE: 24-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/272,796
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J J
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0321 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-855-0555
;             TELEFAX: 415-845-4166
;             TELEX: <Unknown>
;     INFORMATION FOR SEQ ID NO: 15:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 607 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GenBank
;             CLONE: 1827450
;        SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-769-970-15

  Query Match           15.4%;  Score 290.5;  DB 12;  Length 607;
  Best Local Similarity   29.2%;  Pred. No. 6.7e-17;
  Matches   92;  Conservative   45;  Mismatches  129;  Indels   49;  Gaps   12;

Qy         48 LPLSPP----TAPDRATAVATASRLGPYVLLEPEEGG--RAYRALHCPTGTEYTCKVYPV 101
Db          5 LPTSDPGRLITDPRSGRTYLKGRLLG--------KGGPARCYEATDTETGSAYAVKVIPQ 56

Qy        102 QEALAVLEPYARLP-----------PHKHVARPTEVLAGTQLLYAPFTR-THGDMHSLVR 149
Db         57 SR---VAKPHQREKILNEIELHRDLQHRHIVRFSHHFEDADNIYIFLELCSRKSLAHIWK 113

Qy        150 SRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVL 209
Db        114 ARHTLLEPEVRYYLRQILSGLKYLHQRGILHRDLKLGNPFIT--ENMELKVGDFGLAARL 171

Qy        210 TGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPV 267
Db        172 EPPEQR--KKTICGTPNYVAPEVLLRQG--HGPEADVWSLGCVMYTLLCGSPPFETADLK 227

Qy        268 LLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQ----DPMPLA 323
Db        228 ETYRCIKQVHYTLPASLSLPARQLLAAILRASPRDRPSIDQILRHDFFTKGYTPDRLPIS 287

Qy        324 PTRSHLWEAAQVVPD 338
Db        288 --------SCVTVPD 294


RESULT 13
US-10-204-041-16
; Sequence 16, Application US/10204041
; Publication No. US20030176443A1
; GENERAL INFORMATION:
;  APPLICANT: STEIN-GERLACH, MATTHIAS
;  APPLICANT:  SALASSIDIS, KONSTADINOS
;  APPLICANT:  BACHER, GERALD
;  APPLICANT:  MULLER, STEFAN
;  TITLE OF INVENTION: Pyridylpyrimidine Derivatives as Effective Compounds Against Pr
;  TITLE OF INVENTION:  Infections and Prion Diseases
;  FILE REFERENCE: AXM-007.1P US
;  CURRENT APPLICATION NUMBER: US/10/204,041
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: EP 01111858.5
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: PCT/EP02/05420
;  PRIOR FILING DATE: 2002-05-16
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
;   LENGTH: 607
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-204-041-16

  Query Match           15.4%;  Score 290.5;  DB 12;  Length 607;
  Best Local Similarity   29.2%;  Pred. No. 6.7e-17;
  Matches   92;  Conservative   45;  Mismatches  129;  Indels   49;  Gaps   12;

Qy         48 LPLSPP----TAPDRATAVATASRLGPYVLLEPEEGG--RAYRALHCPTGTEYTCKVYPV 101
Db          5 LPTSDPGRLITDPRSGRTYLKGRLLG--------KGGPARCYEATDTETGSAYAVKVIPQ 56

Qy        102 QEALAVLEPYARLP-----------PHKHVARPTEVLAGTQLLYAPFTR-THGDMHSLVR 149
Db         57 SR---VAKPHQREKILNEIELHRDLQHRHIVRFSHHFEDADNIYIFLELCSRKSLAHIWK 113

Qy        150 SRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVL 209
Db        114 ARHTLLEPEVRYYLRQILSGLKYLHQRGILHRDLKLGNPFIT--ENMELKVGDFGLAARL 171

Qy        210 TGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPV 267
Db        172 EPPEQR--KKTICGTPNYVAPEVLLRQG--HGPEADVWSLGCVMYTLLCGSPPFETADLK 227

Qy        268 LLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQ----DPMPLA 323
Db        228 ETYRCIKQVHYTLPASLSLPARQLLAAILRASPRDRPSIDQILRHDFFTKGYTPDRLPIS 287

Qy        324 PTRSHLWEAAQVVPD 338
Db        288 --------SCVTVPD 294
```

```
RESULT 14
US-10-108-580-2
; Sequence 2, Application US/10108580
; Publication No. US20030077681A1
; GENERAL INFORMATION:
;  APPLICANT: Cogswell, John
;  TITLE OF INVENTION: PLK3 PROTEIN-PROTEIN INTERACTIONS
;  FILE REFERENCE: PU4458
;  CURRENT APPLICATION NUMBER: US/10/108,580
;  CURRENT FILING DATE:  2002-03-28
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 607
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-108-580-2

  Query Match             15.4%;  Score 290.5;  DB 15;  Length 607;
  Best Local Similarity   29.2%;  Pred. No. 6.7e-17;
  Matches   92;  Conservative   45;  Mismatches  129;  Indels   49;  Gaps   12;

Qy         48 LPLSPP----TAPDRATAVATASRLGPYVLLEPEEGG--RAYRALHCPTGTEYTCKVYPV 101
              || |     | |          ||          :||  | | |      ||: |  || |
Db          5 LPTSDPGRLITDPRSGRTYLKGRLLG--------KGGFARCYEATDTETGSAYAVKVIPQ 56

Qy        102 QEALAVLEPYARLP-----------PHKHVARPTEVLAGTQLLYAFFTR-THGDMHSLVR 149
                     | :|: |            |:|: | :           :| |     :    :  : :
Db         57 SR---VAKPHQREKILNEIELHRDLQHRHIVRFSHHFEDADNIYIFLELCSRKSLAHIWK 113

Qy        150 SRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVL 209
              :|| : |||      ||: : | : || |:: ||||  |       | :| :  :    :  |
Db        114 ARHTLLEPEVRYYLRQILSGLKYLHQRGILHRDLKLGNFFIT--ENMELKVGDFGLAARL 171

Qy        210 TGPDDSLWDKHAC--PAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPV 267
                |:       |  |  | || ||:|   :    |  ||||||  ::|:|  |  ||: ::
Db        172 EPPEQR--KKTICGTPNYVAPEVLLRQG--HGPEADVWSLGCVMYTLLCGSPPFETADLK 227

Qy        268 LLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQ----DPMPLA 323
                :  |::  | |||  ||  |||  |:  :||   | :| :    || | :   :    | :|::
Db        228 ETYRCIKQVHYTLPASLSLPARQLLAAILRASPRDRPSIDQILRHDFFTKGYTPDRLPIS 287

Qy        324 PTRSHLWEAAQVVPD 338
                         :  |||
Db        288 --------SCVTVPD 294

RESULT 15
US-09-801-368-152
; Sequence 152, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250Alman, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 152
;   LENGTH: 1518
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-152

  Query Match             15.2%;  Score 287;  DB 10;  Length 1518;
  Best Local Similarity   26.7%;  Pred. No. 4.5e-16;
  Matches   86;  Conservative   50;  Mismatches  128;  Indels   58;  Gaps    9;

Qy         56 PDRATAVATAS-------RLGPYVL---LEPEEGGRAYRALHCPTGTEYTCKVYPVQEAL 105
              ||   :|||  |         :||: |    |     ||    | :   ||       |: | ::|
Db         59 PDSTVSVATKSSKRKSRDTVGPWKLGKTLGKGSSGRVRLAKNMETGQLAAIKIVPKKKAF 118

Qy        106 ----------------------------------AVLEPYA--------RLPPHKHVAR 122
                                                  :  ||          :|  | :|
Db        119 VHCSNNGTVPNSYSSSMVTSNVSSPSIASREHSNHSQTNPYGIEREIVIMKLISHTNVMA 178

Qy        123 PTEVLAGTQLLYAFFTRTH-GDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLR 181
                ||     ||        |::   : |: ::|| ||   |:|:    :::||   :  |
Db        179 LFEVWENKSELYLVLEYVDGGELFDYLVSKGKLPEREAIHYFKQIVEGVSYCHSFNICHR 238

Qy        182 DLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKA 241
              |||      :   |:: ::: : :          |   |:  |          | |   |||:  |   | |
Db        239 DLK-PENLLLDKKNRRIKIADF-GMAALELPNKLLKTSCGSPHYASPEIVMGR-PYHGGP 295

Qy        242 ADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPA 301
              :|||| |: || :| || || |    |  |:: | | :|: ||: || |:  :|  :|
Db        296 SDVWSCGIVLFALLTGHLPFNDDNIKKLLLKVQSGKYQMPSNLSSEARDLISKILVIDPE 355

Qy        302 ERLTATGILLHPWLRQ-DPMPL 322
              :|:|   || ||  ::: | :|:
Db        356 KRITTQEILKHPLIKKYDDLPV 377

Search completed: January 15, 2004, 15:03:53
Job time : 40 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:55:46 ; Search time 44 Seconds
(without alignments)
1291.457 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 1891
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1887	99.8	358	22	AAU03509	Human protein kina
2	1884	99.6	358	22	AAM38908	Human polypeptide
3	1884	99.6	358	22	AAB20326	Human protein phos
4	1884	99.6	393	22	AAM40694	Human polypeptide
5	1870	98.9	360	23	ABB80976	Human tribbles hom
6	1675	88.6	360	21	AAY69157	PeptideA JJ503-KS
7	1582.5	83.7	323	22	AAB85791	Human kinase PKIN-
8	1102	58.3	233	21	AAB43657	Human cancer assoc
9	1097	58.0	278	23	ABB06093	Human NS protein s
10	1097	58.0	278	23	ABB06108	Human NS protein s
11	744.5	39.4	372	23	ABB80975	Human tribbles hom
12	681.5	36.0	269	22	AAU28135	Novel human secret
13	659.5	34.9	290	22	AAU28323	Novel human secret
14	578.5	30.6	206	24	ABP96856	Human C8FW protein
15	435	23.0	484	22	ABB71379	Drosophila melanog
16	408	21.6	138	22	ABG58479	Human liver peptid
17	408	21.6	138	22	ABB43079	Peptide #10585 enc
18	408	21.6	138	22	AAM63978	Human brain expres
19	408	21.6	138	22	AAM76798	Human bone marrow
20	408	21.6	138	22	AAM36904	Peptide #10941 enc
21	408	21.6	138	23	ABG45957	Human peptide enco
22	354	18.7	153	21	AAB43922	Human cancer assoc
23	316.5	16.7	371	21	AAG54418	Zea mays protein f
24	297.5	15.7	778	24	ABB98743	Human kinase #1.
25	297	15.7	327	21	AAG54419	Zea mays protein f
26	295	15.6	794	22	AAU03517	Human protein kina
27	295	15.6	794	23	AAE16271	Human kinase PKIN-
28	291.5	15.4	605	22	ABG16826	Novel human diagno
29	291	15.4	523	21	AAB03425	Wheat putative car
30	291	15.4	523	24	ABR40719	Triticum aestivum
31	290.5	15.4	607	24	AAE34495	Human PRK protein.
32	287.5	15.2	754	24	AAO16604	Human cell cycle-r
33	287.5	15.2	1078	24	ABP96069	Human protein kina
34	287	15.2	512	23	AAM50578	Arabidopsis SNF-1
35	284	15.0	504	24	ABR40815	Cucumis sativus oi
36	284	15.0	512	21	AAG36157	Arabidopsis thalia
37	283.5	15.0	619	20	ABB04768	Human dominant neg
38	283.5	15.0	643	20	ABB04769	Mouse dominant neg
39	283.5	15.0	651	23	AAB47857	KIAA0175 protein.
40	283.5	15.0	651	24	ABU57635	Differentially exp
41	283.5	15.0	651	24	ABU56727	Lung cancer-associ
42	283	15.0	438	21	AAB03421	Soybean putative c
43	283	15.0	438	24	ABR40715	Glycine max oil tr
44	283	15.0	512	23	AAO17663	A thaliana AKIN11.
45	283	15.0	512	23	AAO17664	A thaliana AKIN11

ALIGNMENTS

RESULT 1
AAU03509
ID AAU03509 standard; Protein; 358 AA.
XX
AC AAU03509;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human protein kinase #9.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX
DR WPI; 2001-343950/36.
DR N-PSDB; AAS06709.

```
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections -
XX
PS   Claim 7; Figure 2; 433pp; English.
XX
CC   AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC   novel protein kinases have been identified as members of the tyrosine
CC   or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be
CC   used for gene therapy and as DNA probes in diagnostic assays.
CC   The protein kinase polypeptides may be used as antigens in the production
CC   of antibodies against the protein kinases and in assays to identify
CC   modulators of protein kinase expression and activity.
XX
SQ   Sequence   358 AA;

  Query Match             99.8%;  Score 1887;  DB 22;  Length 358;
  Best Local Similarity   99.7%;  Pred. No. 2e-171;
  Matches  357;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db         61 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120

Qy        121 ARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300

Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358


RESULT 2
AAM38908
ID   AAM38908 standard; Protein; 358 AA.
XX
AC   AAM38908;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2053.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58064.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 3; SEQ ID NO 2053; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   358 AA;

  Query Match             99.6%;  Score 1884;  DB 22;  Length 358;
  Best Local Similarity   99.4%;  Pred. No. 3.9e-171;
  Matches  356;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||||:||||||||||:||||||||||||||||||||||||||
Db         61 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCRVYPVQEALAVLEPYARLPPHKHV 120

Qy        121 ARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
```

```
Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARBEEGDREVVLYG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARBEEGDREVVLYG 358


RESULT 3
AAB20326
ID   AAB20326 standard; Protein; 358 AA.
XX
AC   AAB20326;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Human protein phosphatase and kinase protein-5.
XX
KW   Protein phosphatase and kinase protein; PPHKP-5; human;
KW   gastrointestinal disorder; immune system disorder;
KW   neurological disorder; cell proliferative disorder; cancer;
KW   diagnosis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          142..315
FT                   /note= "protein kinase signature"
FT   Domain          242..264
FT                   /note= "tyrosine kinase catalytic domain signature"
FT   Domain          78..303
FT                   /note= "protein kinase domain signature"
FT   Modified-site   54
FT                   /note= "O-phosphorylated"
FT   Modified-site   90
FT                   /note= "O-phosphorylated"
FT   Modified-site   140
FT                   /note= "O-phosphorylated"
FT   Modified-site   210
FT                   /note= "O-phosphorylated"
FT   Modified-site   215
FT                   /note= "O-phosphorylated"
FT   Modified-site   14
FT                   /note= "O-phosphorylated"
FT   Modified-site   27
FT                   /note= "O-phosphorylated"
FT   Modified-site   95
FT                   /note= "O-phosphorylated"
FT   Modified-site   232
FT                   /note= "O-phosphorylated"
FT   Modified-site   238
FT                   /note= "O-phosphorylated"
XX
PN   WO200120004-A2.
XX
PD   22-MAR-2001.
XX
PF   14-SEP-2000; 2000WO-US25515.
XX
PR   15-SEP-1999;   99US-0154141.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Tang YT,  Bandman O,  Hillman JL,  Baughn MR,  Azimzai Y;
PI   Lu DAM;
XX
DR   WPI; 2001-244811/25.
DR   N-PSDB; AAF30480.
XX
PT   Novel human protein phosphatase and kinase proteins for diagnosis,
PT   treatment and prevention of gastrointestinal, immune system,
PT   neurological and cell proliferative disorders -
XX
PS   Claim 1; Page 88-89; 103pp; English.
XX
CC   The present sequence is that of novel human protein phosphatase
CC   and kinase protein PPHKP-5, as predicted from Incyte Clone ID No.
CC   1271505CB1 (see AAF30480). Tissues that express PPHKP-5 (as a
CC   fraction of total tissues expressing PPHKP-5) include reproductive
CC   (0.288), gastrointestinal (0.212) and haematopoietic or immune
CC   (0.192). Diseases or conditions associated with tissues expressing
CC   PPHKP-5 (as a fraction of total tissues expressing PPHKP-5) include
CC   cancer (0.577), inflammation or trauma (0.327) and cell
CC   proliferation (0.308). The encoded protein shows homology to rat
CC   kinase. The invention provides human PPHKP-1 to -11 polypeptides
CC   (see AAB20322-32) and polynucleotides (see AAF30476-86). It also
CC   provides expression vectors, host cells, antibodies, agonists and
CC   antagonists, as well as methods for diagnosing, treating or
CC   preventing disorders associated with expression of PPHKP, including
CC   gastrointestinal disorders, immune system disorders, neurological
CC   disorders and cell proliferative disorders, including cancer.
XX
SQ   Sequence   358 AA;

  Query Match             99.6%;  Score 1884;  DB 22;  Length 358;
  Best Local Similarity   99.4%;  Pred. No. 3.9e-171;
  Matches  356;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db         61 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRTRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300

Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARBEEGDREVVLYG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARBEEGDREVVLYG 358


RESULT 4
AAM40694
ID   AAM40694 standard; Protein; 393 AA.
XX
AC   AAM40694;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 5625.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
```

```
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI59850.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 2; SEQ ID NO 5625; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   393 AA;

  Query Match             99.6%;  Score 1884;  DB 22;  Length 393;
  Best Local Similarity   99.4%;  Pred. No. 4.4e-171;
  Matches  356;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 95

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||:|||||||||||:||||||||||||||||||||||||||||
Db         96 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCRVYPVQEALAVLEPYARLPPHKHV 155

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 215

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 275

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 335

Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 393


RESULT 5
ABB80976
ID   ABB80976 standard; Protein; 360 AA.
XX
AC   ABB80976;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human tribbles homologue-3 (htrb-3) polypeptide encoding DNA.
XX
KW   Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic; osteopathic;
KW   opthalamological; cardiant; cytostatic; haemostatic; immunosuppressive;
KW   antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
KW   tumour necrosis factor; TNF; htrb-3.
XX
OS   Homo sapiens.
XX
PN   WO200253743-A2.
XX
PD   11-JUL-2002.
XX
PF   08-JAN-2002; 2002WO-US00070.
XX
PR   08-JAN-2001; 2001US-260294P.
XX
PA   (INTE-) INTERLEUKIN GENETICS INC.
XX
PI   Dower S,  Quanstrom E,  Kiss-Toth E;
XX
DR   WPI; 2002-590635/63.
DR   N-PSDB; ABN86479.
XX
PT   Novel isolated human tribbles homologue-1 polypeptide for inhibiting
PT   AP-1-mediated inflammatory signal in a cell, and activating
PT   ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
PT   -
XX
PS   Example 10; Fig 11B; 131pp; English.
XX
CC   The invention provides an isolated human tribbles homologue-1 (htrb-1,
CC   also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC   htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC   signal in a cell. The polypeptide employed in the method is preferably
CC   htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or
CC   htrb-3 N C. It is also useful for providing htrb agonist activity for
CC   activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC   signal, an estrogen receptor-mediated gene activation signal, an
CC   fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
CC   in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC   inflammatory signal in a cell such as tumor necrosis factor (TNF)
CC   induced inflammatory signal, or an interleukin induced inflammatory
CC   signal. htrb proteins are useful in screening assays, predictive medicine
CC   and in therapeutics or prophylactics. The htrb proteins are useful for
CC   screening compounds e.g. for treating and/or preventing diseases caused
CC   by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
CC   psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
CC   and cancers. The htrb therapeutics are useful for antagonizing
CC   interleukin-1 dependent disorders of human placenta, intraventricular
CC   hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
CC   and inflammation or autoimmune disorders. The present sequence represents
CC   the htrb-3 polypeptide.
XX
SQ   Sequence   360 AA;

  Query Match             98.9%;  Score 1870;  DB 23;  Length 360;
  Best Local Similarity   98.6%;  Pred. No. 8.5e-170;
  Matches  355;  Conservative    3;  Mismatches    0;  Indels    2;  Gaps    1;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||:|||||||||||:||||||||||||||||||||||||||||
```

```
Db         61 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARVPPHKHV 120

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180

Qy        181 RDLKLCRFVFA--DRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYS 238
              |||||||||||  |||:||||||||||||||||||||||||||||||||||||||||||||
Db        181 RDLKLCRFVFADRDREKKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYS 240

Qy        239 GKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRR 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRR 300

Qy        299 EPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 360
```

RESULT 6
AAY69157
ID AAY69157 standard; Peptide; 360 AA.
XX
AC AAY69157;
XX
DT 30-MAY-2000 (first entry)
XX
DE PeptideA JJ503-KS comprising domains VIA to XI of a protein kinase.
XX
KW Kinase activity; molecular weight marker; isoelectric focusing marker;
KW peptide fragmentation control; cellular signal transduction.
XX
OS Homo sapiens.
XX
PN WO200008180-A2.
XX
PD 17-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17630.
XX
PR 04-AUG-1998; 98US-0095270.
PR 11-SEP-1998; 98US-0099972.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Virca GD, Bird TA, Anderson DM, Marken JS;
XX
DR WPI; 2000-195584/17.
DR N-PSDB; AAZ61155.
XX
PT New human kinase polypeptides and polynucleotides used as molecular
PT weight markers and as controls for peptide fragmentation -
XX
PS Claim 2; Page 10; 60pp; English.
XX
CC The present sequence represents a partial polypeptide which has kinase
CC activity. The kinase polynucleotides can be used to express the
CC polypeptides, and as probes to identify nucleic acids encoding
CC proteins having kinase activity. The kinase polypeptides and
CC fragmented polypeptides are used as molecular weight and isoelectric
CC focusing markers, and as controls for peptide fragmentation. They also
CC have a number of therapeutic uses as kinases play a central role in
CC cellular signal transduction. The polypeptides could also be used to
CC identify binding partner proteins. The polypeptides can also be used as
CC a reagent to identify any proteins that the polypeptide regulates, and
CC proteins with which it might interact. The polypeptides may also be
CC used for preparation of antibodies. The antibodies can be used in
CC assays to detect the presence of the protein, and to purify the protein
CC by immunoaffinity chromatography.
XX
SQ Sequence 360 AA;

Query Match 88.6%; Score 1675; DB 21; Length 360;
Best Local Similarity 99.7%; Pred. No. 3.5e-151;
Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 102

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        103 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 162

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 222

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 282

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 342

Qy        301 AERLTATGILLHPWLRQD 318
              ||||||||||||||||||
Db        343 AERLTATGILLHPWLRQD 360
```

RESULT 7
AAB85791
ID AAB85791 standard; Protein; 323 AA.
XX
AC AAB85791;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human kinase PKIN-10.
XX
KW PKIN; kinase; cytostatic; immunosuppressive; immunostimulant; human;
KW antiarteriosclerotic; cardiant; gene therapy; antisense therapy.
XX
OS Homo sapiens.
XX
PN WO200160991-A2.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US05240.
XX
PR 17-FEB-2000; 2000US-0183682.
PR 02-MAR-2000; 2000US-0186559.
PR 09-MAR-2000; 2000US-0188606.
PR 17-MAR-2000; 2000US-0189998.
PR 30-MAR-2000; 2000US-0193851.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Tang YT, Buford N, Gandhi AR, Patterson C, Khan FA, Yue H;
PI Hafalia A, Shih LL, Tribouley CM, Yao MG, Burrill JD, Marcus GA;
PI Zingler KA, Lu DAM, Bandman O, Policky JL, Griffin JA, Thornton M;
PI Nguyen DB, Lal P, Walsh RT;
XX
DR WPI; 2001-514771/56.
DR N-PSDB; AAH76218.
XX
PT isolated human kinase polypeptides useful in the diagnosis, treatment
PT and prevention of cancer, immune disorders and disorders affecting
PT growth and development -
XX
PS Claim 1; Page 115; 126pp; English.
XX
CC The invention provides human kinases (PKIN) and polynucleotides encoding
CC PKIN. The PKIN polypeptides can be expressed using standard recombinant

```
CC   methodology. The PKIN polypeptides, polynucleotides, modulators and
CC   specific antibodies are useful in the diagnosis, treatment and prevention
CC   of cancer, immune disorders, disorders affecting growth and development,
CC   atherosclerosis, and other cardiovascular diseases, and lipid disorders
CC   and in the assessment of the effects of exogenous compounds on the
CC   expression of nucleic acid sequences of human kinases. The present
CC   sequence represents a human PKIN-10 polypeptide.
XX
SQ   Sequence   323 AA;

  Query Match             83.7%;  Score 1582.5;  DB 22;  Length 323;
  Best Local Similarity   86.0%;  Pred. No. 2e-142;
  Matches  308;  Conservative    4;  Mismatches   11;  Indels   35;  Gaps    2;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRATPLAASAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              |||||||||||||||||||||:||||||||||||||||||||||||||||||| |||
Db         61 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPDKHV 120

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              | ||| |||||||||||| |||||| |:                  |  |||
Db        121 AWPTEGLAGTQLLYAFFTWTHGDMHRLIGH-----------------TPCAHC------- 156

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 240
                  |: ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 ----------DQTRKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGK 205

Qy        241 AADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 300
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AANVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREP 265

Qy        301 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARREEGDREVVLYG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 AERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEARREEGDREVVLYG 323


RESULT 8
AAB43657
ID   AAB43657 standard; Protein; 233 AA.
XX
AC   AAB43657;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated protein sequence SEQ ID NO:1102.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   N-PSDB; AAC77866.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 11; Page 1707-1708; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence   233 AA;

  Query Match             58.3%;  Score 1102;  DB 21;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-97;
  Matches  208;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        151 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 85

Qy        211 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 145

Qy        271 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 205

Qy        331 EAAQVVPDGLGLDEAREEEGDREVVLYG 358
              ||||||||||||||||||||||||||||
Db        206 EAAQVVPDGLGLDEAREEEGDREVVLYG 233


RESULT 9
ABB06093
ID   ABB06093 standard; Protein; 278 AA.
XX
AC   ABB06093;
XX
DT   10-MAY-2002  (first entry)
XX
DE   Human NS protein sequence SEQ ID NO:185.
XX
KW   Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW   antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW   vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW   anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW   antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW   anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW   gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW   contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW   endometriosis; degenerative disease; multiple sclerosis; psoriasis;
```

KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO200206315-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-IL00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Freilich S, Bernstein J;
XX
DR WPI; 2002-155037/20.
DR N-PSDB; ABL39747.
XX
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 6; Page 213-214; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SQ Sequence 278 AA;

Query Match 58.0%; Score 1097; DB 23; Length 278;
Best Local Similarity 99.1%; Pred. No. 3.5e-96;
Matches 209; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 88

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         89 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 148

Qy        121 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        149 ARPTEVLAGTQLLYAFFTRTHGDMHSLVRSPHRIPEPEAAVLFRQMATALAHCHQHGLVL 208

Qy        181 RDLKLCRFVFADRERKKLVLENLEDSCVLTG 211
              |||||||||||||||||||||||||||||||
Db        209 RDLKLCRFVFADRERKKLVLENLEDSCVLTG 239
```

RESULT 10
ABB06108
ID ABB06108 standard; Protein; 278 AA.
XX
AC ABB06108;
XX
DT 10-MAY-2002 (first entry)
XX
DE Human NS protein sequence SEQ ID NO:200.
XX
KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO200206315-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-IL00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Freilich S, Bernstein J;
XX
DR WPI; 2002-155037/20.
DR N-PSDB; ABL39762.
XX
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 6; Page 231-232; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SQ Sequence 278 AA;

```
  Query Match             58.0%;  Score 1097;  DB 23;  Length 278;
  Best Local Similarity   99.1%;  Pred. No. 3.5e-96;
  Matches  209;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 88

Qy         61 AVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 120
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db         89 AVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHV 148

Qy        121 ARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVL 180
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        149 ARPTEVLAGTQLLYAPFTRTHGDMHSLVRSPHRIPEPEAAVLFRQMATALAHCHQHGLVL 208

Qy        181 RDLKLCRFVPADRERKKLVLENLEDSCVLTG 211
              |||||||||||||||||||||||||||||||
Db        209 RDLKLCRFVPADRERKKLVLENLEDSCVLTG 239


RESULT 11
ABB80975
ID   ABB80975 standard; Protein; 372 AA.
XX
AC   ABB80975;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human tribbles homologue-1 (htrb-1) polypeptide.
XX
KW   Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic; osteopathic;
KW   opthalamological; cardiant; cytostatic; haemostatic; immunosuppressive;
KW   antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
KW   tumour necrosis factor; TNF.
XX
OS   Homo sapiens.
XX
PN   WO200253743-A2.
XX
PD   11-JUL-2002.
XX
PF   08-JAN-2002; 2002WO-US00070.
XX
PR   08-JAN-2001; 2001US-260294P.
XX
PA   (INTE-) INTERLEUKIN GENETICS INC.
XX
PI   Dower S,  Quanstrom E,  Kiss-Toth E;
XX
DR   WPI; 2002-590635/63.
DR   N-PSDB; ABN86478.
XX
PT   Novel isolated human tribbles homologue-1 polypeptide for inhibiting
PT   AP-1-mediated inflammatory signal in a cell, and activating
PT   ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
PT   -
XX
PS   Claim 21; Fig 10B; 131pp; English.
XX
CC   The invention provides an isolated human tribbles homologue-1 (htrb-1,
CC   also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC   htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC   signal in a cell. The polypeptide employed in the method is preferably
CC   htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or
CC   htrb-3 N C. It is also useful for providing htrb agonist activity for
CC   activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC   signal, an estrogen receptor-mediated gene activation signal, an
CC   fibroblast growth factor (FGP) induced signal, or a PMA induced signal,
CC   in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC   inflammatory signal in a cell such as tumor necrosis factor (TNP)
CC   induced inflammatory signal, or an interleukin induced inflammatory
CC   signal. htrb proteins are useful in screening assays, predictive medicine
CC   and in therapeutics or prophylactics. The htrb proteins are useful for
CC   screening compounds e.g. for treating and/or preventing diseases caused
CC   by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
CC   psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
CC   and cancers. The htrb therapeutics are useful for antagonizing
CC   interleukin-1 dependent disorders of human placenta, intraventricular
CC   hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
CC   and inflammation or autoimmune disorders. The present sequence represents
CC   the htrb-1 polypeptide.
XX
SQ   Sequence   372 AA;

  Query Match             39.4%;  Score 744.5;  DB 23;  Length 372;
  Best Local Similarity   46.8%;  Pred. No. 2.3e-62;
  Matches  162;  Conservative   47;  Mismatches  126;  Indels   11;  Gaps    6;

Qy          2 RATPLAAPA-GSLSRKKRLELDD--NLDTERPVQKRARSGP---QPRLPPCLLPLSPPTA 55
              |   |  ||   :  |: |: ||   :  : |      | |    |   ||  |  || |
Db         17 RGPALLFPATRGVPAKRLLDADDAAAVAAKCPRLSECSSPPDYLSPPGSPC-SPQPPPAA 75

Qy         56 PDRATAVATA---SRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQEALAVLEPYA 112
              |       :|   ||:  |:||   |     |||   || |  |||:|::     : ||
Db         76 PGAGGGSGSAPGPSRIADYLLLPLAEREHVSRALCIHTGRELRCKVFPIKHYQDKIRPYI 135

Qy        113 RLPPHKHVARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAH 172
              :|| | ::    ||: |    | || :: ||||| |||| |: | ||| ||:|: :|:||
Db        136 QLPSHSNITGIVEVILGETKAYVFFEKSFGDMHSYVRSRKRLREEEAARLFKQIVSAVAH 195

Qy        173 CHQHGLVLRDLKLCRFVPADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILS 232
              |||  :|| |||| :|||:  || :| ||:|||: :: | ||:| ||| ||||| ||||:
Db        196 CHQSAIVLGDLKLRKFVFSTEERTQLRLESLEDTHIMKGEDDALSDKHGCPAYVSPEILN 255

Qy        233 SRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLV 292
              :  :|||||||||||| |:|:| | ||| ||:|  || |||||  : :|  :|  ||||:
Db        256 TTGTYSGKAADVWSLGVMLYTLLVGRYPFHDSDPSALFSKIRRGQFCIPEHISPKARCLI 315

Qy        293 RCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPD 338
              | ||||||:|||||  ||||||      |     | :  : |:||:
Db        316 RSLLRREPSERLTAPEILLHPWFESVLEP-GYIDSEIGTSDQIVPE 360


RESULT 12
AAU28135
ID   AAU28135 standard; Protein; 269 AA.
XX
AC   AAU28135;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secretory protein, Seq ID No 304.
XX
KW   Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW   ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW   transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW   amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW   ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW   gut protection; lung; liver fibrosis; immune deficiency; infection;
KW   severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW   multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW   fertility; analgesic; pain; antigen.
XX
OS   Homo sapiens.
XX
PN   WO200166689-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US04942.
XX
PR   07-MAR-2000; 2000US-0519705.
PR   19-MAY-2000; 2000US-0574454.
PR   17-JUN-2000; 2000US-0596193.
```

```
PR   14-JUL-2000; 2000US-0616847.
PR   19-SEP-2000; 2000US-0665363.
PR   20-OCT-2000; 2000US-0693267.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Xu C,  Wehrman T,  Ren F,  Ma Y,  Zhou P;
PI   Zhao QA,  Yang Y,  Drmanac RT,  Zhang J,  Chen R,  Xue AJ,  Wang J;
XX
DR   WPI; 2001-589934/66.
DR   N-PSDB; AAS45035.
XX
PT   Novel polypeptides and nucleic acids obtained from cDNA libraries
PT   prepared from various human tissues, for diagnosis and treatment of
PT   cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS   Example 4; SEQ ID No 304; 107pp; English.
XX
CC   The invention relates to novel isolated human secreted polypeptides (I)
CC   and polynucleotides (II). (I) and (II) are useful for treating
CC   inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC   ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC   involved in increasing haematopoiesis, stem cell survival, bone growth
CC   and remodeling. (I), (II) and modulators of (II) are useful for
CC   prophylaxis or treatment of one or more cancers. (II) is also useful for
CC   creating transgenic animals useful for studying the in vivo activities of
CC   the polypeptide as well as for studying modulators of the polypeptides.
CC   (I) induces the proliferation of neural cells and regeneration of nerve
CC   and brain tissue and is useful for the treatment of central and
CC   peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC   Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC   sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC   activity, regulation of haematopoiesis and is useful for treating myeloid
CC   or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC   and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC   tissue growth, and in tissue repair, healing of burns, incisions,
CC   ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC   disorders, or periodontal disease. Furthermore, (I) is also useful for
CC   gut protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues, various immune deficiencies and
CC   disorders including severe combined immunodeficiency (SCID), bacterial or
CC   fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC   rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC   reactions and conditions, such as asthma or other respiratory problems.
CC   In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC   fertility, metabolism, catabolism, anabolism, storage or elimination of
CC   dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC   analgesic effects or other pain reducing effects, immunoglobulin like
CC   activity and can act as an antigen in a vaccine composition to raise an
CC   immune response. AAU28020-AAU28395 represent novel human secreted protein
CC   amino acid sequences of the invention.
XX
SQ   Sequence   269 AA;

  Query Match             36.0%;  Score 681.5;  DB 22;  Length 269;
  Best Local Similarity   51.8%;  Pred. No. 1.5e-56;
  Matches  133;  Conservative   38;  Mismatches   79;  Indels    7;  Gaps    2;

Qy         88 CP------TGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTH 141
              ||      |    |||:|::    : || :|| | ::    ||: |    | || :
Db          2 CPGRCASTLGRRVRCKVFPIKHYQDKIRPYIQLPSHSNITGIVEVILGETKAYVFFEKDF 61

Qy       142 GDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLE 201
             ||||| |||| |: | ||| ||:|: :|:|||||  :|| |||| :|||:  || :| ||
Db        62 GDMHSYVRSRKRLREEEAARLFKQIVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLE 121

Qy       202 NLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPF 261
             :|||: :: | ||:| ||| ||||| ||||::  :|||||||||||||| |:|:| | |||
Db       122 SLEDTHIMKGEDDALSDKHGCPAYVSPEILNTTGTYSGKAADVWSLGVMLYTLLVGRYPF 181

Qy       262 QDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMP 321
              ||:|  || ||||| : :|  :|  ||||:| ||||||:|||||  ||||||      |
Db       182 HDSDPSALFSKIRRGQFCIPEHISPKARCLIRSLLRREPSERLTAPEILLHPWFESVLEP 241

Qy       322 LAPTRSHLWEAAQVVPD 338
                 | :  : |:||:
Db       242 -GYIDSEIGTSDQIVPE 257


RESULT 13
AAU28323
ID   AAU28323 standard; Protein; 290 AA.
XX
AC   AAU28323;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secretory protein, Seq ID No 680.
XX
KW   Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW   ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW   transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW   amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW   ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW   gut protection; lung; liver fibrosis; immune deficiency; infection;
KW   severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW   multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW   fertility; analgesic; pain; antigen.
XX
OS   Homo sapiens.
XX
PN   WO200166689-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US04942.
XX
PR   07-MAR-2000; 2000US-0519705.
PR   19-MAY-2000; 2000US-0574454.
PR   17-JUN-2000; 2000US-0596193.
PR   14-JUL-2000; 2000US-0616847.
PR   19-SEP-2000; 2000US-0665363.
PR   20-OCT-2000; 2000US-0693267.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Xu C,  Wehrman T,  Ren F,  Ma Y,  Zhou P;
PI   Zhao QA,  Yang Y,  Drmanac RT,  Zhang J,  Chen R,  Xue AJ,  Wang J;
XX
DR   WPI; 2001-589934/66.
DR   N-PSDB; AAS45223.
XX
PT   Novel polypeptides and nucleic acids obtained from cDNA libraries
PT   prepared from various human tissues, for diagnosis and treatment of
PT   cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS   Example 2; SEQ ID No 680; 107pp; English.
XX
CC   The invention relates to novel isolated human secreted polypeptides (I)
CC   and polynucleotides (II). (I) and (II) are useful for treating
CC   inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC   ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC   involved in increasing haematopoiesis, stem cell survival, bone growth
CC   and remodeling. (I), (II) and modulators of (II) are useful for
CC   prophylaxis or treatment of one or more cancers. (II) is also useful for
CC   creating transgenic animals useful for studying the in vivo activities of
CC   the polypeptide as well as for studying modulators of the polypeptides.
CC   (I) induces the proliferation of neural cells and regeneration of nerve
CC   and brain tissue and is useful for the treatment of central and
CC   peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC   Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC   sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC   activity, regulation of haematopoiesis and is useful for treating myeloid
CC   or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC   and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC   tissue growth, and in tissue repair, healing of burns, incisions,
```

CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention.
XX
SQ Sequence 290 AA;

Query Match 34.9%; Score 659.5; DB 22; Length 290;
Best Local Similarity 51.2%; Pred. No. 2.1e-54;
Matches 127; Conservative 42; Mismatches 78; Indels 1; Gaps 1;

```
Qy         91 GTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRS 150
Db         22 GPPFSLQVFPIKHYQDKIRPYIQLPSHSNITGIVEVILGETKAYVFFEKDFGDMHSYVRS 81

Qy        151 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 210
Db         82 RKRLREEEAARLFKQIVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLESLEDTHIMK 141

Qy        211 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 270
Db        142 GEDDALSDKHGCPAYVSPEILNTTGTYSGKAADVWTGGVMLYTLLVGRYPFHDSDPSALF 201

Qy        271 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 330
Db        202 SKIRRGQFCIPEHISPKARCLIRSLLRREPSERLTAPEILLHPWFESVLEP-GYIDSEIG 260

Qy        331 EAAQVVPD 338
Db        261 TSDQIVPE 268
```

RESULT 14
ABP96856
ID ABP96856 standard; Protein; 206 AA.
XX
AC ABP96856;
XX
DT 11-JUN-2003 (first entry)
XX
DE Human C8FW protein SEQ ID NO:2.
XX
KW Human; phosphoprotein regulating mitogenic pathway gene; cytostatic;
KW vaccine; cancer; C8FW.
XX
OS Homo sapiens.
XX
PN WO2003018769-A2.
XX
PD 06-MAR-2003.
XX
PF 27-AUG-2002; 2002WO-US27187.
XX
PR 27-AUG-2001; 2001US-314655P.
PR 31-OCT-2001; 2001US-330797P.
XX
PA (TULA-) TULARIK INC.
XX
PI Li J;
XX
DR WPI; 2003-278653/27.
DR N-PSDB; ACC45125, ACC45126.
XX
PT Diagnosing, preventing, treating and prognosticating a cancer in a
PT mammal by detecting and measuring the C8FW gene copy number and/or C8FW
PT level, useful for breast, brain, lung, colon, ovarian and/or prostate
PT cancers -
XX
PS Claim 16; Page 86; 86pp; English.
XX
CC The present invention describes a method for diagnosing a cancer in a
CC mammal. The method comprises detecting and measuring the C8FW gene copy
CC number or level in a biological subject from a region of the mammal that
CC is suspected to be pre-cancerous or cancerous to generate data for a
CC test gene copy, and comparing the test gene copy number or level to a
CC data for a control gene copy number or level, where an amplification of
CC the gene in the biological subject relative to the control indicates the
CC presence of a pre-cancerous lesion or cancer in the animal. C8FW has
CC cytostatic activity, and can be used in vaccines. The method can be used
CC for the diagnosis, prevention, treatment and prognostication of breast,
CC colon, lung, brain, prostate and/or ovarian cancer. The present sequence
CC represents human C8FW from the present invention. C8FW is a
CC phosphoprotein regulating mitogenic pathway protein.
XX
SQ Sequence 206 AA;

Query Match 30.6%; Score 578.5; DB 24; Length 206;
Best Local Similarity 57.9%; Pred. No. 7.1e-47;
Matches 113; Conservative 29; Mismatches 52; Indels 1; Gaps 1;

```
Qy        144 MHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENL 203
Db          1 MHSYVRSRKRLREEEAARLFKQIVSAVAHCHQSAIVLGDLKLRKFVFSTEERTQLRLESL 60

Qy        204 EDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQD 263
Db         61 EDTHIMKGEDDALSDKHGCPAYVSPEILNTTGTYSGKAADVWSLGVMLYTLLVGRYPFHD 120

Qy        264 SEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLA 323
Db        121 SDPSALFSKIRRGQFCIPEHISPKARCLIRSLLRREPSERLTAPEILLHPWFESVLEP-G 179

Qy        324 PTRSHLWEAAQVVPD 338
Db        180 YIDSEIGTSDQIVPE 194
```

RESULT 15
ABB71379
ID ABB71379 standard; Protein; 484 AA.
XX
AC ABB71379;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 40929.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX

DR WPI; 2001-656860/75.
DR N-PSDB; ABL15482.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 40929; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 484 AA;

Query Match 23.0%; Score 435; DB 22; Length 484;
Best Local Similarity 33.4%; Pred. No. 1.1e-32;
Matches 116; Conservative 45; Mismatches 134; Indels 52; Gaps 8;

```
Qy          31 VQKRARSGPQPRLPPCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYRALHCPT 90
               :|:|     ||      :    | |  |     ||| |                | :   |
Db         104 IQQRYLISAQPSHISAAVAAKTP-ASYRHLVDLTASNL--------------RCVDIFT 147

Qy          91 GTEYTCKVYPVQEAL-AVLEPYARLPPHKHVARPTEVLA---------------GTQLL 133
               | :: |::  | | |  |   | :|  |     | : :                 | :|
Db         148 GEQFLCRI--VNEPLHKVQRAYFQLQQHDEELRRSTIYGHPLIRPVHDIIPLTKDRTYIL 205

Qy         134 YA----------FFTRTHGDMHSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDL 183
                |           |  : ::|: :|   |: | ||  :| |:   :  ||::|::||||
Db         206 IAPVPQERDSTGGVTGVYENLHTYIRHAKRLCETEARAIFHQICQTVQVCHRNGIILRDL 265

Qy         184 KLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAAD 243
               || || | |  | ||  |:|| | :| | ||:| ||  || |  ||:|  : :| || ||
Db         266 KLKRFYFIDEARTKLQYESLEGSMILDGEDDTLSDKIGCPLYTAPELLCPQQTYKGKPAD 325

Qy         244 VWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAER 303
               :|||| |:||| | ||| :      |   || |   :|  ||   | |:  |||::  ||
Db         326 MWSLGVILYTMLVGQYPFYEKANCNLITVIRHGNVQIPLTLSKSVRWLLLSLLRKDYTER 385

Qy         304 LTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEG 350
               :||: | | ||||:      |   :|    :|  |     :| |:||
Db         386 MTASHIFLTPWLREQ----RPFHMYLPVDVEVAED---WSDAEEDEG 425
```

Search completed: January 15, 2004, 14:57:50
Job time : 46 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:03:07 ; Search time 41 Seconds
(without alignments)
2253.240 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 358
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14	3.9	257	11	Q8R2V8	Q8r2v8 mus musculu
2	14	3.9	343	4	Q92519	Q92519 homo sapien
3	14	3.9	343	6	Q28283	Q28283 canis famil
4	14	3.9	343	11	Q8K4K3	Q8k4k3 mus musculu
5	14	3.9	343	11	Q8K017	Q8k017 mus musculu
6	13	3.6	224	4	O15180	O15180 homo sapien
7	13	3.6	364	11	Q9EQL6	Q9eql6 rattus norv
8	13	3.6	372	4	Q96RU8	Q96ru8 homo sapien
9	13	3.6	372	4	Q9H2Y8	Q9h2y8 homo sapien
10	13	3.6	372	11	Q8K4K4	Q8k4k4 mus musculu
11	13	3.6	372	11	Q91W04	Q91w04 mus musculu
12	10	2.8	339	10	Q8H7S4	Q8h7s4 oryza sativ
13	9	2.5	317	16	Q9F3C2	Q9f3c2 streptomyce
14	9	2.5	562	5	Q18041	Q18041 caenorhabdi
15	9	2.5	583	5	Q9XZ66	Q9xz66 caenorhabdi
16	8	2.2	184	10	Q9LRS3	Q9lrs3 arabidopsis
17	8	2.2	190	16	Q92XU5	Q92xu5 rhizobium m
18	8	2.2	245	2	Q9EZ07	Q9ez07 streptomyce
19	8	2.2	249	16	Q9L184	Q9l184 streptomyce
20	8	2.2	277	4	Q9NS63	Q9ns63 homo sapien
21	8	2.2	325	11	Q9JJD6	Q9jjd6 mus musculu
22	8	2.2	325	11	Q8BNY7	Q8bny7 mus musculu
23	8	2.2	367	10	Q9AQV3	Q9aqv3 oryza sativ
24	8	2.2	375	16	Q8F930	Q8f930 leptospira
25	8	2.2	376	4	Q9BYG4	Q9byg4 homo sapien
26	8	2.2	427	9	Q8SD32	Q8sd32 pseudomonas
27	8	2.2	428	16	Q9RDL4	Q9rdl4 streptomyce
28	8	2.2	499	11	Q8C0V7	Q8c0v7 mus musculu
29	8	2.2	575	4	Q9H6R5	Q9h6r5 homo sapien
30	8	2.2	575	4	Q8TBX5	Q8tbx5 homo sapien
31	8	2.2	608	12	O41253	O41253 rice ragged
32	8	2.2	696	4	Q8IVF5	Q8ivf5 homo sapien
33	8	2.2	725	16	Q9HZG0	Q9hzg0 pseudomonas
34	8	2.2	813	16	Q8YJM9	Q8yjm9 brucella me
35	8	2.2	814	16	Q8FY69	Q8fy69 brucella su
36	8	2.2	934	6	Q9GMD3	Q9gmd3 bos taurus
37	8	2.2	1159	5	Q9VVC9	Q9vvc9 drosophila
38	8	2.2	1715	11	Q9WVS3	Q9wvs3 mus musculu
39	8	2.2	2038	12	Q9IH80	Q9ih80 cherry necr
40	7	2.0	59	10	Q8GY77	Q8gy77 arabidopsis
41	7	2.0	66	16	Q9Z4Y0	Q9z4y0 streptomyce
42	7	2.0	72	16	Q8U8M5	Q8u8m5 agrobacteri
43	7	2.0	76	11	Q8K4U2	Q8k4u2 mus musculu
44	7	2.0	80	16	Q99SP0	Q99sp0 staphylococ
45	7	2.0	88	12	Q8QWS9	Q8qws9 norwalk vir

ALIGNMENTS

RESULT 1
Q8R2V8
ID Q8R2V8 PRELIMINARY; PRT; 257 AA.
AC Q8R2V8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 29.2 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC027159; AAH27159.1; -.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; ATP-binding; Transferase.
FT NON_TER 1 1
SQ SEQUENCE 257 AA; 29232 MW; 84EE4C0DD4767F51 CRC64;

Query Match 3.9%; Score 14; DB 11; Length 257;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 236 SYSGKAADVWSLGV 249
||||||||||||||
Db 143 SYSGKAADVWSLGV 156

RESULT 2
Q92519
ID Q92519 PRELIMINARY; PRT; 343 AA.
AC Q92519;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)

```
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   GS3955 (GS3955 protein).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cancellous bone;
RA   Ohno I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Uterus;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D87119; BAA13250.1; -.
DR   EMBL; BC002637; AAH02637.1; -.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   343 AA;  38800 MW;  BF8B7366DACB84FA CRC64;

  Query Match             3.9%;  Score 14;  DB 4;  Length 343;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        236 SYSGKAADVWSLGV 249
              ||||||||||||||
Db        229 SYSGKAADVWSLGV 242


RESULT 3
Q28283
ID   Q28283      PRELIMINARY;      PRT;   343 AA.
AC   Q28283;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   C5FW ORF protein.
GN   C5FW ORF.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RX   MEDLINE=97067069; PubMed=8910471;
RA   Wilkin F., Savonet V., Radulescu A., Petermans J., Dumont J.E.,
RA   Maenhaut C.;
RT   "Identification and Characterization of Novel Genes Modulated in the
RT   Thyroid of Dogs Treated with Methimazole and Propylthiouracil.";
RL   J. Biol. Chem. 271:28451-28457(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RX   MEDLINE=98000262; PubMed=9342215;
RA   Wilkin F., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,
RA   Dumont J.E., Maenhaut C.;
RT   "Characterization of a phosphoprotein whose mRNA is regulated by the
RT   mitogenic pathways in dog thyroid cells.";
RL   Eur. J. Biochem. 248:660-669(1997).
DR   EMBL; X99144; CAA67581.1; -.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   343 AA;  38786 MW;  BF8D1300DACB84FA CRC64;

  Query Match             3.9%;  Score 14;  DB 6;  Length 343;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        236 SYSGKAADVWSLGV 249
              ||||||||||||||
Db        229 SYSGKAADVWSLGV 242


RESULT 4
Q8K4K3
ID   Q8K4K3      PRELIMINARY;      PRT;   343 AA.
AC   Q8K4K3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   TRB-2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kiss-Toth E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA   Bagstaff S.M., Wyllie D.H., Harte M., O'Neill L.A.J., Qwarnstrom E.E.,
RA   Dower S.K.;
RT   "Mammalian homologs of Drosophila tribbles (htrb) control mitogen
RT   activated protein kinase signaling.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF358867; AAM45477.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   343 AA;  38758 MW;  0B3965B8B2087D74 CRC64;

  Query Match             3.9%;  Score 14;  DB 11;  Length 343;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        236 SYSGKAADVWSLGV 249
              ||||||||||||||
Db        229 SYSGKAADVWSLGV 242


RESULT 5
Q8K017
ID   Q8K017      PRELIMINARY;      PRT;   343 AA.
AC   Q8K017;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   TRB-2 (C5FW ORF protein homolog).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Aorta and vein, Cerebellum, and Retina;
```

```
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; BC034338; AAH34338.1; -.
DR   EMBL; AK044747; BAC32063.1; -.
DR   EMBL; AK080064; BAC37820.1; -.
DR   EMBL; AK082329; BAC38467.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   343 AA;  38772 MW;  9418B7AC19FCC23F CRC64;

  Query Match             3.9%;  Score 14;  DB 11;  Length 343;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        236 SYSGKAADVWSLGV 249
              ||||||||||||||
Db        229 SYSGKAADVWSLGV 242


RESULT 6
O15180
ID   O15180      PRELIMINARY;      PRT;   224 AA.
AC   O15180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Phosphoprotein (Fragment).
GN   C8FW.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RA   Wilkain F.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid;
RX   MEDLINE=98000262; PubMed=9342215;
RA   Wilkin F., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,
RA   Dumont J.E., Maenhaut C.;
RT   "Characterization of a phosphoprotein whose mRNA is regulated by the
RT   mitogenic pathways in dog thyroid cells.";
RL   Eur. J. Biochem. 248:660-669(1997).
DR   EMBL; AJ000480; CAA04119.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
FT   NON_TER       1       1
SQ   SEQUENCE   224 AA;  25479 MW;  740CD0905F86499B CRC64;

  Query Match             3.6%;  Score 13;  DB 4;  Length 224;
  Best Local Similarity   100.0%;  Pred. No. 0.00043;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVWSLGV 249
              |||||||||||||
Db        112 YSGKAADVWSLGV 124


RESULT 7
Q9EQL6
ID   Q9EQL6      PRELIMINARY;      PRT;   364 AA.
AC   Q9EQL6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   G-protein-coupled receptor induced protein GIG2 (Fragment).
GN   GIG2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RA   Mayhaus M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,
RA   Nitsch R.M.;
RT   "Identification of a novel nuclear factor Gig2, as an m1-acetylcholine
RT   receptor-induced gene.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF205438; AAG35664.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Receptor; Transferase.
FT   NON_TER       1       1
SQ   SEQUENCE   364 AA;  40377 MW;  8F9167FB76DFCD37 CRC64;

  Query Match             3.6%;  Score 13;  DB 11;  Length 364;
  Best Local Similarity   100.0%;  Pred. No. 0.00066;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVWSLGV 249
              |||||||||||||
Db        253 YSGKAADVWSLGV 265


RESULT 8
Q96RU8
ID   Q96RU8      PRELIMINARY;      PRT;   372 AA.
AC   Q96RU8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   SKIP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kiss-Toth E., Wyllie D.H., Qwarnstrom E.E., Dower S.K.;
RT   "Identification of pro-inflammatory cytokine signalling network
RT   components by transcription expression screening.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF250310; AAK58174.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   372 AA;  40980 MW;  3E2B5C87A4F98FDB CRC64;

  Query Match             3.6%;  Score 13;  DB 4;  Length 372;
  Best Local Similarity   100.0%;  Pred. No. 0.00067;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVWSLGV 249
              |||||||||||||
```

```
Db        260 YSGKAADVWSLGV 272


RESULT 9
Q9H2Y8
ID   Q9H2Y8      PRELIMINARY;      PRT;   372 AA.
AC   Q9H2Y8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   G-protein-coupled receptor induced protein GIG2.
GN   GIG2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Mayhaus M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,
RA   Nitsch R.M.;
RT   "Identification of a novel nuclear factor Gig2, as an m1-acetylcholine
RT   receptor-induced gene.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF205437; AAG35663.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Receptor; Transferase.
SQ   SEQUENCE   372 AA;  41008 MW;  5F54E50924B1365B CRC64;

  Query Match             3.6%;  Score 13;  DB 4;  Length 372;
  Best Local Similarity  100.0%;  Pred. No. 0.00067;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVWSLGV 249
              |||||||||||||
Db        260 YSGKAADVWSLGV 272


RESULT 10
Q8K4K4
ID   Q8K4K4      PRELIMINARY;      PRT;   372 AA.
AC   Q8K4K4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   TRB-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kiss-Toth E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA   Bagstaff S.M., Wyllie D.H., Harte M., O'Neill L.A.J., Qwarnstrom E.E.,
RA   Dower S.K.;
RT   "Mammalian homologs of Drosophila tribbles (htrb) control mitogen
RT   activated protein kinase signaling.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF358866; AAM45478.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   372 AA;  41282 MW;  3A3DE82B46CD907F CRC64;

  Query Match             3.6%;  Score 13;  DB 11;  Length 372;
  Best Local Similarity  100.0%;  Pred. No. 0.00067;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVWSLGV 249
              |||||||||||||
Db        260 YSGKAADVWSLGV 272


RESULT 11
Q91W04
ID   Q91W04      PRELIMINARY;      PRT;   372 AA.
AC   Q91W04;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to phosphoprotein regulated by mitogenic pathways.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Skin;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; BC006800; AAH06800.1; -.
DR   EMBL; AK028626; BAC26038.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Transferase.
SQ   SEQUENCE   372 AA;  41281 MW;  AD29BB4E640B4B62 CRC64;

  Query Match             3.6%;  Score 13;  DB 11;  Length 372;
  Best Local Similarity  100.0%;  Pred. No. 0.00067;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVWSLGV 249
              |||||||||||||
Db        260 YSGKAADVWSLGV 272


RESULT 12
Q8H7S4
ID   Q8H7S4      PRELIMINARY;      PRT;   339 AA.
AC   Q8H7S4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   OSJNBA0081P02.7.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,
RA   Currie J., Collura K.;
RT   "Rice Genomic Sequence.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC107226; AAN52742.1; -.
KW   Hypothetical protein.
```

```
SQ   SEQUENCE   339 AA;  36383 MW;  4514B42089AB607B CRC64;

  Query Match             2.8%;  Score 10;  DB 10;  Length 339;
  Best Local Similarity   100.0%;  Pred. No. 0.52;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 AADVWSLGVA 250
              ||||||||||
Db        235 AADVWSLGVA 244


RESULT 13
Q9F3C2
ID   Q9F3C2      PRELIMINARY;      PRT;   317 AA.
AC   Q9F3C2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative oxidoreductase.
GN   SCO7553 OR SC5F1.07.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Saunders D.C., Harris D.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939132; CAC16433.1; -.
DR   InterPro; IPR002085; Adh_zn_family.
DR   Pfam; PF00107; adh_zinc; 1.
KW   Complete proteome.
SQ   SEQUENCE   317 AA;  31455 MW;  5AB9335A378352CD CRC64;

  Query Match             2.5%;  Score 9;  DB 16;  Length 317;
  Best Local Similarity   100.0%;  Pred. No. 4.6;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        280 LPAGLSAPA 288
              |||||||||
Db        107 LPAGLSAPA 115


RESULT 14
Q18041
ID   Q18041      PRELIMINARY;      PRT;   562 AA.
AC   Q18041;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 63.4 kDa protein.
GN   C16B8.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bentley D.;
RT   "The sequence of C. elegans cosmid C16B8.";
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U41031; AAA82618.2; -.
DR   HSSP; P11362; 1FGK.
DR   WormPep; C16B8.1; CE27692.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   InterPro; IPR003306; WIF.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF02019; WIF; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00469; WIF; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   Hypothetical protein; ATP-binding; Kinase; Transferase;
KW   Tyrosine-protein kinase.
SQ   SEQUENCE   562 AA;  63406 MW;  F366B69F7F84C916 CRC64;

  Query Match             2.5%;  Score 9;  DB 5;  Length 562;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 AADVWSLGV 249
              |||||||||
Db        477 AADVWSLGV 485


RESULT 15
Q9XZ66
ID   Q9XZ66      PRELIMINARY;      PRT;   583 AA.
AC   Q9XZ66;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Tyrosine kinase receptor-related protein RYK precursor.
GN   RYK.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99167506; PubMed=10066802;
RA   Halford M.M., Oates A.C., Hibbs M.L., Wilks A.F., Stacker S.A.;
RT   "Genomic structure and expression of the mouse growth factor receptor
RT   related to tyrosine kinases (Ryk).";
RL   J. Biol. Chem. 274:7379-7390(1999).
DR   EMBL; AF133217; AAD24877.1; -.
DR   HSSP; P11362; 1FGK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   InterPro; IPR003306; WIF.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF02019; WIF; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00469; WIF; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   ATP-binding; Kinase; Receptor; Signal; Transferase;
KW   Tyrosine-protein kinase.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    583       TYROSINE KINASE RECEPTOR-RELATED PROTEIN.
SQ   SEQUENCE   583 AA;  65901 MW;  FE6DFF8E47D24300 CRC64;

  Query Match             2.5%;  Score 9;  DB 5;  Length 583;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 AADVWSLGV 249
              |||||||||
Db        498 AADVWSLGV 506


Search completed: January 15, 2004, 15:06:48
Job time : 42 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:01:57 ; Search time 17 Seconds
(without alignments)
990.327 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 358
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	274	76.5	358	1	NIPK_HUMAN	Q96ru7 homo sapien
2	25	7.0	349	1	NIPK_RAT	Q9wtq6 rattus norv
3	22	6.1	354	1	NIPK_MOUSE	Q8k4k2 mus musculu
4	8	2.2	174	1	HUNB_DROMU	O46250 drosophila
5	8	2.2	376	1	PA6G_HUMAN	Q9byg4 homo sapien
6	8	2.2	431	1	NRKA_TRYBB	Q08942 trypanosoma
7	8	2.2	431	1	NRKB_TRYBB	Q03428 trypanosoma
8	7	2.0	82	1	YKPA_YEAST	O60200 saccharomyc
9	7	2.0	91	1	E310_ADE03	P11318 human adeno
10	7	2.0	91	1	E310_ADE07	P15134 human adeno
11	7	2.0	116	1	MERT_PSEAE	P04140 pseudomonas
12	7	2.0	156	1	ATPF_VIBAL	P12989 vibrio algi
13	7	2.0	195	1	TATB_XANCP	Q8p3h9 xanthomonas
14	7	2.0	201	1	CTF1_HUMAN	Q16619 homo sapien
15	7	2.0	204	1	PAAD_BACSU	P94404 bacillus su
16	7	2.0	209	1	YQIJ_ECOLI	P76657 escherichia
17	7	2.0	219	1	RS19_AQUAE	O66435 aquifex aeo
18	7	2.0	246	1	ADC_CLOBE	Q9rpk1 clostridium
19	7	2.0	295	1	NMCR_ENTCL	P52676 enterobacte
20	7	2.0	315	1	OMPP_ECOLI	P34210 escherichia
21	7	2.0	316	1	L767_CAEEL	Q09517 caenorhabdi
22	7	2.0	324	1	HEM3_PAEMA	O69110 paenibacill
23	7	2.0	332	1	LPXK_PSEAE	Q9hzm3 pseudomonas
24	7	2.0	344	1	RLA0_AERPE	Q9y9w8 aeropyrum p
25	7	2.0	349	1	NRLB_KLEPO	P10045 klebsiella
26	7	2.0	370	1	MYOM_APLCA	P15513 a myomoduli
27	7	2.0	378	1	LUXE_VIBHA	P14286 vibrio harv
28	7	2.0	397	1	DXR_HAEIN	P44055 haemophilus
29	7	2.0	419	1	CPEY_SYNPY	Q02174 synechococc
30	7	2.0	419	1	MPK7_HUMAN	O14733 homo sapien
31	7	2.0	435	1	FUZ7_USTMA	Q99078 ustilago ma
32	7	2.0	441	1	NFIX_MESAU	P13623 mesocricetu
33	7	2.0	451	1	CBPS_STRGR	P18143 streptomyce
34	7	2.0	460	1	VATB_METAC	Q8tij0 methanosarc
35	7	2.0	460	1	VATB_METBA	P22663 methanosarc
36	7	2.0	460	1	VATB_METMA	Q60187 methanosarc
37	7	2.0	487	1	HEP_DROME	Q23977 drosophila
38	7	2.0	507	1	GTR6_HUMAN	Q9ugq3 homo sapien
39	7	2.0	515	1	AMPA_MYCTU	Q10401 mycobacteri
40	7	2.0	519	1	GLG2_SOLTU	P55242 solanum tub
41	7	2.0	535	1	TX21_HUMAN	Q9ul17 homo sapien
42	7	2.0	545	1	MUTL_THETH	Q9ra54 thermus the
43	7	2.0	595	1	SYD_BACHD	Q9kdg1 bacillus ha
44	7	2.0	605	1	WIS1_SCHPO	P33886 schizosacch
45	7	2.0	614	1	MGM_EUBBA	Q59268 eubacterium

ALIGNMENTS

RESULT 1
NIPK_HUMAN
ID NIPK_HUMAN STANDARD; PRT; 358 AA.
AC Q96RU7; Q8TAI6; Q9H5M8; Q9NUD2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neuronal cell death inducible putative kinase (SKIP3).
GN NIPK OR C20ORF97.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiss-Toth E., Wyllie D.H., Qwarnstrom E.E., Dower S.K.;
RT "Identification of pro-inflammatory cytokine signalling network
RT components by transcription expression screening.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).

```
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cervix, and Muscle;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May play an important role in a common pathway leading
CC       to programmed neuronal cell death. Does not appear to function in
CC       the programmed death of non-neuronal cells. May serve as an
CC       endogenous antagonist competing for substrate with functional
CC       kinases that act to promote neuronal cell survival (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 protein kinase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF250311; AAK58175.1; -.
DR   EMBL; AK026945; BAB15597.1; -.
DR   EMBL; AL034548; CAB81634.1; -.
DR   EMBL; BC019363; AAH19363.1; -.
DR   EMBL; BC027484; AAH27484.1; -.
DR   Genew; HGNC:16228; C20orf97.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Apoptosis.
FT   DOMAIN       68    316       PROTEIN KINASE.
FT   CONFLICT     84     84       Q -> R (IN REF. 4; AAH27484).
FT   CONFLICT    105    105       L -> P (IN REF. 2).
FT   CONFLICT    114    114       L -> V (IN REF. 1).
FT   CONFLICT    194    195       ER -> DREK (IN REF. 1).
SQ   SEQUENCE   358 AA;  39577 MW;  CE15FD89A81E8D63 CRC64;

  Query Match             76.5%;  Score 274;  DB 1;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-272;
  Matches  274;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDM 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDM 144

Qy        145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVPADRERKKLVLENLE 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVPADRERKKLVLENLE 204

Qy        205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 264

Qy        265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAP 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAP 324

Qy        325 TRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              ||||||||||||||||||||||||||||||||||
Db        325 TRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358


RESULT 2
NIPK_RAT
ID   NIPK_RAT       STANDARD;      PRT;   349 AA.
AC   Q9WTQ6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Neuronal cell death inducible putative kinase.
GN   NIPK.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Neuron;
RX   MEDLINE=99262087; PubMed=10329375;
RA   Mayumi-Matsuda K., Kojima S., Suzuki H., Sakata T.;
RT   "Identification of a novel kinase-like gene induced during neuronal
RT   cell death.";
RL   Biochem. Biophys. Res. Commun. 258:260-264(1999).
CC   -!- FUNCTION: May play an important role in a common pathway leading
CC       to programmed neuronal cell death. Does not appear to function
CC       in the programmed death of non-neuronal cells. May serve as an
CC       endogenous antagonist competing for substrate with functional
CC       kinases that act to promote neuronal cell survival.
CC   -!- TISSUE SPECIFICITY: Detected only in the lung. Not detected in the
CC       heart, brain, spleen, liver, skeletal muscle, kidney and testis.
CC   -!- INDUCTION: Expression induced during programmed cell death evoked
CC       in neuronal cells by NGF-depletion.
CC   -!- SIMILARITY: Contains 1 protein kinase domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB020967; BAA77582.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Apoptosis.
FT   DOMAIN       63    310       PROTEIN KINASE.
SQ   SEQUENCE   349 AA;  38602 MW;  3050F9BFE946D815 CRC64;

  Query Match              7.0%;  Score 25;  DB 1;  Length 349;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-17;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        210 TGPDDSLWDKHACPAYVGPEILSSR 234
              |||||||||||||||||||||||||
Db        205 TGPDDSLWDKHACPAYVGPEILSSR 229


RESULT 3
NIPK_MOUSE
ID   NIPK_MOUSE     STANDARD;      PRT;   354 AA.
AC   Q8K4K2; Q921E7;
```

```
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Neuronal cell death inducible putative kinase (TRB-3).
GN   NIPK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kiss-Toth E., Dempsey C., Jozsa V., Caunt J., Oxley K.M.,
RA   Bagstaff S.M., Wyllie D.H., Harte M., O'Neill L.A.J., Qwarnstrom E.E.,
RA   Dower S.K.;
RT   "Mammalian homologs of Drosophila tribbles (htrb) control mitogen
RT   activated protein kinase signaling.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May play an important role in a common pathway leading
CC       to programmed neuronal cell death. Does not appear to function in
CC       the programmed death of non-neuronal cells. May serve as an
CC       endogenous antagonist competing for substrate with functional
CC       kinases that act to promote neuronal cell survival (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 protein kinase domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF358868; AAM45476.1; -.
DR   EMBL; AK089931; BAC41002.1; -.
DR   EMBL; BC012955; AAH12955.1; -.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Apoptosis.
FT   DOMAIN       68    315       PROTEIN KINASE.
FT   CONFLICT    157    157       S -> P (IN REF. 3).
FT   CONFLICT    219    219       K -> T (IN REF. 1).
FT   CONFLICT    239    264       MISSING (IN REF. 3).
FT   CONFLICT    301    354       SERLVALGILLHPWLREDHGRVSPPQSDRREMDQVVPDGPQ
FT                                LEEAEEGEVGLYG -> CRATCGPGNPLASLVERGSRPSLS
FT                                STV (IN REF. 2).
SQ   SEQUENCE   354 AA;  39022 MW;  2CB283FC119F859P CRC64;

  Query Match             6.1%;  Score 22;  DB 1;  Length 354;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-14;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        213 DDSLWDKHACPAYVGPEILSSR 234
              ||||||||||||||||||||||
Db        213 DDSLWDKHACPAYVGPEILSSR 234


RESULT 4
HUNB_DROMU
ID   HUNB_DROMU     STANDARD;      PRT;   174 AA.
AC   O46250; O46251;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Hunchback protein (Fragments).
GN   HB.
OS   Drosophila mulleri (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7231;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Baker R.H., Desalle R.;
RT   "Multiple sources of character information and the phylogeny of
RT   Hawaiian Drosophilids.";
RL   Syst. Biol. 46:654-673(1997).
CC   -!- FUNCTION: GAP CLASS SEGMENTATION PROTEIN THAT CONTROLS DEVELOPMENT
CC       OF HEAD STRUCTURES (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE HUNCHBACK FAMILY OF C2H2-TYPE ZINC-
CC       FINGER PROTEINS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U93014; AAC03262.1; -.
DR   EMBL; U93015; AAC03263.1; -.
DR   FlyBase; FBgn0023760; Dmul\hb.
KW   Developmental protein; Gap protein; Zinc-finger;
KW   Metal-binding; DNA-binding; Repeat; Nuclear protein.
FT   NON_TER       1      1
FT   DOMAIN       58     67       POLY-GLN.
FT   NON_CONS     93     94
FT   NON_TER     174    174
SQ   SEQUENCE   174 AA;  19265 MW;  DA27CF6CC8CC8368 CRC64;

  Query Match             2.2%;  Score 8;  DB 1;  Length 174;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         58 RATAVATA 65
              ||||||||
Db        118 RATAVATA 125


RESULT 5
PA6G_HUMAN
ID   PA6G_HUMAN     STANDARD;      PRT;   376 AA.
AC   Q9BYG4;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Partitioning defective-6 homolog gamma (PAR-6 gamma) (PAR6D).
GN   PARD6G OR PAR6G.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND INTERACTION WITH RAC1; CDC42; PRKCI AND PRKCZ.
RC   TISSUE=Neuroblastoma;
RX   PubMed=11260256;
RA   Noda Y., Takeya R., Ohno S., Naito S., Ito T., Sumimoto H.;
RT   "Human homologues of the Caenorhabditis elegans cell polarity protein
RT   PAR6 as an adaptor that links the small GTPases Rac and Cdc42 to
RT   atypical protein kinase C.";
RL   Genes Cells 6:107-119(2001).
RN   [2]
RP   INTERACTION WITH ARHQ.
RX   MEDLINE=20394296; PubMed=10934474;
RA   Joberty G., Petersen C., Gao L., Macara I.G.;
RT   "The cell-polarity protein Par6 links Par3 and atypical protein kinase
RT   C to Cdc42.";
RL   Nat. Cell Biol. 2:531-539(2000).
CC   -!- FUNCTION: Adapter protein involved in asymetrical cell division
CC       and cell polarization processes. May play a role in the formation
CC       of epithelial tight junctions. The PARD6-PARD3 complex links GTP-
CC       bound Rho small GTPases to atypical protein kinase C proteins (By
CC       similarity).
CC   -!- SUBUNIT: Interacts with PARD3 (Probable). Interacts with GTP-bound
CC       forms of CDC42, ARHQ/TC10 and RAC1. Interacts with the N-terminal
CC       part of PRKCI and PRKCZ.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- TISSUE SPECIFICITY: Widely expressed, with a higher expression in
CC       fetal and adult kidney.
CC   -!- DOMAIN: The pseudo-CRIB domain together with the PDZ domain is
CC       required for the interaction with Rho small GTPases (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the PAR6 family.
CC   -!- SIMILARITY: Contains 1 pseudo-CRIB domain.
CC   -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB044556; BAB40757.1; -.
DR   Genew; HGNC:16076; PARD6G.
DR   InterPro; IPR001478; PDZ.
DR   Pfam; PF00595; PDZ; 1.
DR   SMART; SM00666; PB1; 1.
DR   SMART; SM00228; PDZ; 1.
DR   PROSITE; PS50106; PDZ; 1.
KW   Cell cycle; Cell division; Tight junction; Membrane.
FT   DOMAIN      134    151       PSEUDO-CRIB.
FT   DOMAIN      158    251       PDZ.
FT   DOMAIN      127    254       INTERACTION WITH PARD3 AND CDC42
FT                                (BY SIMILARITY).
SQ   SEQUENCE   376 AA;  40883 MW;  E5A4EF521403CA11 CRC64;

  Query Match             2.2%;  Score 8;  DB 1;  Length 376;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 APAGSLSR 15
              ||||||||
Db        318 APAGSLSR 325


RESULT 6
NRKA_TRYBB
ID   NRKA_TRYBB     STANDARD;      PRT;   431 AA.
AC   Q08942;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative serine/threonine-protein kinase A (EC 2.7.1.37).
GN   NRKA.
OS   Trypanosoma brucei brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EATRO 164 / Isolate ISTAR1;
RX   MEDLINE=93295429; PubMed=8515773;
RA   Gale M.J. Jr., Parsons M.;
RT   "A Trypanosoma brucei gene family encoding protein kinases with
RT   catalytic domains structurally related to Nek1 and NIMA.";
RL   Mol. Biochem. Parasitol. 59:111-122(1993).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC   -!- SIMILARITY: Contains 1 PH domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L03778; AAB59252.1; -.
DR   PIR; T11854; T11854.
DR   HSSP; Q63450; 1A06.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00233; PH; 1.
```

```
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding.
FT   DOMAIN       20    279       PROTEIN KINASE.
FT   DOMAIN      331    429       PH.
FT   NP_BIND      26     34       ATP (BY SIMILARITY).
FT   BINDING      49     49       ATP (BY SIMILARITY).
FT   ACT_SITE    147    147       BY SIMILARITY.
FT   VARIANT     192    192       A -> P (IN STRAIN TREU66).
FT   VARIANT     199    199       K -> L (IN STRAIN TREU66).
SQ   SEQUENCE   431 AA;  47915 MW;  2E68B70275884224 CRC64;

  Query Match             2.2%;  Score 8;  DB 1;  Length 431;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        242 ADVWSLGV 249
              ||||||||
Db        205 ADVWSLGV 212


RESULT 7
NRKB_TRYBB
ID   NRKB_TRYBB     STANDARD;      PRT;   431 AA.
AC   Q03428;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative serine/threonine-protein kinase B (EC 2.7.1.37).
GN   NRKB.
OS   Trypanosoma brucei brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate TREU66;
RX   MEDLINE=93295429; PubMed=8515773;
RA   Gale M.J. Jr., Parsons M.;
RT   "A Trypanosoma brucei gene family encoding protein kinases with
RT   catalytic domains structurally related to Nek1 and NIMA.";
RL   Mol. Biochem. Parasitol. 59:111-122(1993).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC   -!- SIMILARITY: Contains 1 PH domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L03777; AAB59253.1; -.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00233; PH; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding.
FT   DOMAIN       20    279       PROTEIN KINASE.
FT   DOMAIN      331    429       PH.
FT   NP_BIND      26     34       ATP (BY SIMILARITY).
FT   BINDING      49     49       ATP (BY SIMILARITY).
FT   ACT_SITE    147    147       BY SIMILARITY.
SQ   SEQUENCE   431 AA;  48172 MW;  F6B4B3EF9E99FF74 CRC64;

  Query Match             2.2%;  Score 8;  DB 1;  Length 431;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        242 ADVWSLGV 249
              ||||||||
Db        205 ADVWSLGV 212


RESULT 8
YKFA_YEAST
ID   YKFA_YEAST     STANDARD;      PRT;    82 AA.
AC   O60200;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical 9.3 kDa protein in OAR1-CSE4 intergenic region.
GN   YKL053C-A.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94378723; PubMed=8091862;
RA   Rasmussen S.W.;
RT   "Sequence of a 28.6 kb region of yeast chromosome XI includes the
RT   FBA1 and TOA2 genes, an open reading frame (ORF) similar to a
RT   translationally controlled tumour protein, one ORF containing motifs
RT   also found in plant storage proteins and 13 ORFs with weak or no
RT   homology to known proteins.";
RL   Yeast 10:S63-S68(1994).
CC   -!- SIMILARITY: BELONGS TO THE UPF0203 (15E1.1) FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z28054; CAA81891.1; ALT_INIT.
DR   EMBL; Z28052; CAA81889.1; ALT_INIT.
DR   SGD; S0007243; YKL053C-A.
DR   GO; GO:0007005; P:mitochondrion organization and biogenesis; IMP.
DR   Pfam; PF05254; UPF0203; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   82 AA;  9296 MW;  6D8EA7CDD89FB73E CRC64;

  Query Match             2.0%;  Score 7;  DB 1;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        342 LDEAREE 348
              |||||||
Db         62 LDEAREE 68


RESULT 9
E310_ADE03
ID   E310_ADE03     STANDARD;      PRT;    91 AA.
AC   P11318;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
```

```
DE   Early E3B 10.4 kDa protein precursor.
OS   Human adenovirus type 3.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=45659;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87219876; PubMed=3582978;
RA   Signaes C., Akusjaervi G., Pettersson U.;
RT   "Region E3 of human adenoviruses; differences between the oncogenic
RT   adenovirus-3 and the non-oncogenic adenovirus-2.";
RL   Gene 50:173-184(1986).
RN   [2]
RP   IDENTIFICATION OF PROTEIN.
RX   MEDLINE=90112650; PubMed=2296083;
RA   Tollefson A.E., Krajsci P., Yei S., Carlin C.R., Wold W.S.M.;
RT   "A 10,400-molecular-weight membrane protein is coded by region E3 of
RT   adenovirus.";
RL   J. Virol. 64:794-801(1990).
CC   -!- FUNCTION: DOWN-REGULATES THE EGF RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
CC       reticulum (Probable).
CC   -!- SIMILARITY: TO EQUIVALENT PROTEIN IN OTHER ADENOVIRUSES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M15952; AAA42487.1; -.
DR   InterPro; IPR005041; Adeno_E3B.
DR   Pfam; PF03376; Adeno_E3B; 1.
DR   ProDom; PD006617; Adeno_E3B; 1.
KW   Early protein; Transmembrane; Signal; Endoplasmic reticulum.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23     91       E3B PROTEIN.
FT   DOMAIN       23     34       LUMENAL (POTENTIAL).
FT   TRANSMEM     35     60       POTENTIAL.
FT   DOMAIN       61     91       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   91 AA;  10356 MW;  10F5281008DB2C41 CRC64;

  Query Match              2.0%;  Score 7;  DB 1;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 TAVATAS 66
              |||||||
Db         22 TAVATAS 28


RESULT 10
E310_ADE07
ID   E310_ADE07     STANDARD;      PRT;    91 AA.
AC   P15134;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Early E3B 10.4 kDa protein precursor.
OS   Human adenovirus type 7.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=10519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Gomen;
RX   MEDLINE=89073758; PubMed=2849239;
RA   Hong J.S., Mullis K.G., Engler J.A.;
RT   "Characterization of the early region 3 and fiber genes of Ad7.";
RL   Virology 167:545-553(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=87-922;
RA   Kajon A.E., Wadell G.;
RL   Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: DOWN-REGULATES THE EGF RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
CC       reticulum (Probable).
CC   -!- SIMILARITY: TO EQUIVALENT PROTEIN IN OTHER ADENOVIRUSES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M23696; AAA53251.1; -.
DR   EMBL; Z48954; CAA88813.1; -.
DR   PIR; C31830; ERAD27.
DR   PIR; G31830; ERAD77.
DR   InterPro; IPR005041; Adeno_E3B.
DR   Pfam; PF03376; Adeno_E3B; 1.
DR   ProDom; PD006617; Adeno_E3B; 1.
KW   Early protein; Transmembrane; Signal; Endoplasmic reticulum.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23     91       E3B PROTEIN.
FT   DOMAIN       23     34       LUMENAL (POTENTIAL).
FT   TRANSMEM     35     60       POTENTIAL.
FT   DOMAIN       61     91       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   91 AA;  10330 MW;  0C85346008CE2C54 CRC64;

  Query Match              2.0%;  Score 7;  DB 1;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 TAVATAS 66
              |||||||
Db         22 TAVATAS 28


RESULT 11
MERT_PSEAE
ID   MERT_PSEAE     STANDARD;      PRT;    116 AA.
AC   P04140;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Mercuric transport protein (Mercury ion transport protein).
GN   MERT.
OS   Pseudomonas aeruginosa.
OG   Plasmid pVS1.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TRANSPOSON=Tn501;
RX   MEDLINE=85014891; PubMed=6091128;
RA   Misra T.K., Brown N.L., Fritzinger D.C., Pridmore R.D., Barnes W.M.,
RA   Haberstroh L., Silver S.;
RT   "Mercuric ion-resistance operons of plasmid R100 and transposon
RT   Tn501: the beginning of the operon including the regulatory region
RT   and the first two structural genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5975-5979(1984).
CC   -!- FUNCTION: INVOLVED IN MERCURIC TRANSPORT. PASSES A HG(2+) ION
CC       FROM THE PERIPLASMIC MERP PROTEIN TO THE MERCURIC REDUCTASE
CC       (MERA).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z00027; CAA77321.1; -.
DR   EMBL; K02503; AAA27433.1; -.
DR   PIR; A04457; QQPSHT.
DR   InterPro; IPR003457; Transprt_MerT.
DR   Pfam; PF02411; MerT; 1.
KW   Transport; Transposable element; Mercuric resistance; Inner membrane;
KW   Mercury; Plasmid; Transmembrane.
FT   TRANSMEM      16     36       POTENTIAL.
FT   TRANSMEM      44     64       POTENTIAL.
FT   TRANSMEM      94    114       POTENTIAL.
FT   METAL         24     24       HG(2+) (POTENTIAL).
FT   METAL         25     25       HG(2+) (POTENTIAL).
FT   METAL         76     76       HG(2+) (POTENTIAL).
FT   METAL         82     82       HG(2+) (POTENTIAL).
SQ   SEQUENCE   116 AA;  12498 MW;  905B5492AFE08E6C CRC64;

  Query Match             2.0%;  Score 7;  DB 1;  Length 116;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        105 LAVLEPY 111
              |||||||
Db         44 LAVLEPY 50


RESULT 12
ATPF_VIBAL
ID   ATPF_VIBAL     STANDARD;      PRT;   156 AA.
AC   P12989;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase B chain (EC 3.6.3.14).
GN   ATPF OR UNCF.
OS   Vibrio alginolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=663;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=138-2;
RX   MEDLINE=90016889; PubMed=2529481;
RA   Krumholz L.R., Esser U., Simoni R.D.;
RT   "Nucleotide sequence of the unc operon of Vibrio alginolyticus.";
RL   Nucleic Acids Res. 17:7993-7994(1989).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC       CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC       SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC       HAS THREE MAIN SUBUNITS: A, B AND C.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X16050; CAA34177.1; -.
DR   PIR; S06078; S06078.
DR   HSSP; P00859; 1B9U.
DR   InterPro; IPR005864; ATPaseB.
DR   InterPro; IPR002146; ATPsynt_B/B'sub.
DR   Pfam; PF00430; ATP-synt_B; 1.
DR   TIGRFAMs; TIGR01144; ATP_synt_b; 1.
KW   Hydrogen ion transport; Transmembrane; CF(0).
FT   TRANSMEM       2     22       POTENTIAL.
SQ   SEQUENCE   156 AA;  17532 MW;  89057460643D8CF1 CRC64;

  Query Match             2.0%;  Score 7;  DB 1;  Length 156;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        342 LDEAREE 348
              |||||||
Db         87 LDEAREE 93


RESULT 13
TATB_XANCP
ID   TATB_XANCP     STANDARD;      PRT;   195 AA.
AC   Q8P3H9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Sec-independent protein translocase protein tatB homolog.
GN   TATB OR XCC4092.
OS   Xanthomonas campestris (pv. campestris).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=340;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
CC   -!- FUNCTION: Required for correct localization of precursor proteins
CC       bearing signal peptides with the twin arginine conserved motif
CC       S/T-R-R-X-F-L-K. This sec-independent pathway is termed TAT for
CC       twin-arginine translocation system. This system mainly transports
CC       proteins with bound cofactors that require folding prior to export
CC       (By similarity).
CC   -!- SUBCELLULAR LOCATION: Inner-membrane bound (Probable).
CC   -!- SIMILARITY: Belongs to the tatB family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE012532; AAM43313.1; -.
DR   HAMAP; MF_00237; -; 1.
DR   InterPro; IPR003369; MttA_Hcf106.
DR   InterPro; IPR003998; TatB.
DR   Pfam; PF02416; MttA_Hcf106; 1.
DR   PRINTS; PR01506; TATBPROTEIN.
DR   TIGRFAMs; TIGR01410; tatB; 1.
KW   Transport; Protein transport; Translocation; Transmembrane;
KW   Inner membrane; Complete proteome.
FT   TRANSMEM       1     21       POTENTIAL.
```

```
SQ   SEQUENCE   195 AA;  20929 MW;  8DDDDB6F526AD390 CRC64;

  Query Match             2.0%;  Score 7;  DB 1;  Length 195;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 ATPLAAP 9
              |||||||
Db        112 ATPLAAP 118


RESULT 14
CTP1_HUMAN
ID   CTP1_HUMAN     STANDARD;      PRT;   201 AA.
AC   Q16619;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cardiotrophin-1 (CT-1).
GN   CTF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=96429882; PubMed=8833032;
RA   Pennica D., Swanson T.A., Shaw K.J., Kuang W.-J., Gray C.L.,
RA   Beatty B.G., Wood W.I.;
RT   "Human cardiotrophin-1: protein and gene structure, biological and
RT   binding activities, and chromosomal localization.";
RL   Cytokine 8:183-189(1996).
CC   -!- FUNCTION: INDUCES CARDIAC MYOCYTE HYPERTROPHY IN VITRO. BINDS TO
CC       AND ACTIVATES THE LEUKEMIA INHIBITORY FACTOR RECEPTOR (LIF
CC       RECEPTOR)/GP 130 RECEPTOR COMPLEX.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, SKELETAL MUSCLE,
CC       PROSTATE AND OVARY. LOWER LEVELS IN LUNG, KIDNEY, PANCREAS,
CC       THYMUS, TESTIS AND SMALL INTESTINE. LITTLE OR NO EXPRESSION IN
CC       BRAIN, PLACENTA, LIVER, SPLEEN, COLON OR PERIPHERAL BLOOD
CC       LEUKOCYTES.
CC   -!- SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U43033; AAD12173.1; -.
DR   EMBL; U43031; AAD12173.1; JOINED.
DR   EMBL; U43032; AAD12173.1; JOINED.
DR   EMBL; U43030; AAA85229.1; -.
DR   PIR; G02312; G02312.
DR   Genew; HGNC:2499; CTF1.
DR   MIM; 600435; -.
DR   GO; GO:0005576; C:extracellular; TAS.
DR   GO; GO:0005146; F:leukemia inhibitory factor receptor ligand . . .; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007517; P:muscle development; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
KW   Cytokine; Polymorphism.
FT   VARIANT      92     92       A -> T (IN dbSNP:2234933).
FT                                /FTId=VAR_014938.
SQ   SEQUENCE   201 AA;  21227 MW;  0235A7B5745F675F CRC64;

  Query Match             2.0%;  Score 7;  DB 1;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        282 AGLSAPA 288
              |||||||
Db         69 AGLSAPA 75


RESULT 15
PAAD_BACSU
ID   PAAD_BACSU     STANDARD;      PRT;   204 AA.
AC   P94404;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable aromatic acid decarboxylase (EC 4.1.1.-).
GN   YCLB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=97124189; PubMed=8969502;
RA   Yamane K., Kumano M., Kurita K.;
RT   "The 25 degrees-36 degrees region of the Bacillus subtilis chromosome:
RT   determination of the sequence of a 146 kb segment and identification
RT   of 113 genes.";
RL   Microbiology 142:3047-3056(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE /
CC       PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; D50453; BAA08996.1; -.
DR   EMBL; Z99105; CAB12157.1; -.
DR   EMBL; Z99106; CAB12171.1; -.
DR   PIR; G69761; G69761.
DR   SubtiList; BG12023; yclB.
DR   InterPro; IPR003382; Flavoprotein.
DR   InterPro; IPR004507; UbiX.
DR   Pfam; PF02441; Flavoprotein; 1.
DR   TIGRFAMs; TIGR00421; ubiX; 1.
KW   Hypothetical protein; Lyase; Decarboxylase; Complete proteome.
SQ   SEQUENCE   204 AA;  22539 MW;  F6C1A5219365D728 CRC64;

  Query Match             2.0%;  Score 7;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        194 ERKKLVL 200
              |||||||
Db        124 ERKKLVL 130


Search completed: January 15, 2004, 15:05:20
Job time : 18 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:02:32 ; Search time 21 Seconds
(without alignments)
1639.447 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 358
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9	2.5	548	2	T15523	hypothetical prote
2	8	2.2	190	2	A95405	conserved hypothet
3	8	2.2	431	2	T11854	protein kinase (EC
4	8	2.2	725	2	A83266	conserved hypothet
5	8	2.2	813	2	AH3258	cation-transportin
6	7	2.0	33	2	T36211	probable excisiona
7	7	2.0	66	2	H44057	early E3 10.3K pro
8	7	2.0	66	2	T36199	probable excisiona
9	7	2.0	72	2	AD3056	hypothetical prote
10	7	2.0	80	2	H89990	hypothetical prote
11	7	2.0	86	2	S78725	protein YKL053c-a
12	7	2.0	91	1	ERAD27	early E3 10.3K pro
13	7	2.0	91	1	ERAD77	early E3 10.3K pro
14	7	2.0	116	1	QQPSHT	hypothetical prote
15	7	2.0	123	2	B69418	hypothetical prote
16	7	2.0	131	2	E72649	hypothetical prote
17	7	2.0	134	2	D84672	hypothetical prote
18	7	2.0	143	2	G95179	conserved hypothet
19	7	2.0	154	2	AH3143	hypothetical prote
20	7	2.0	156	1	S06078	H+-transporting tw
21	7	2.0	158	2	D98144	hypothetical prote
22	7	2.0	162	2	G98046	conserved hypothet
23	7	2.0	186	2	H70300	ribosomal protein
24	7	2.0	201	2	G02312	cardiotrophin-1 -
25	7	2.0	201	2	T25706	hypothetical prote
26	7	2.0	204	2	G69761	phenylacrylic acid
27	7	2.0	204	2	D91120	probable oxidoredu
28	7	2.0	204	2	D85965	probable oxidoredu
29	7	2.0	204	2	D81793	hypothetical prote
30	7	2.0	205	2	AH2790	conserved hypothet
31	7	2.0	205	2	G97569	BH2278 hypothetica
32	7	2.0	209	2	H65092	hypothetical prote
33	7	2.0	222	2	T31246	hypothetical prote
34	7	2.0	235	2	H75539	branched-chain ami
35	7	2.0	237	2	A75400	probable phosphogl
36	7	2.0	246	2	A84384	hypothetical prote
37	7	2.0	281	2	B83724	transcription regu
38	7	2.0	282	2	I51236	brain-derived neur
39	7	2.0	290	2	B83876	D-alanyl-D-alanine
40	7	2.0	295	2	S35914	regulatory protein
41	7	2.0	295	2	T34774	hypothetical prote
42	7	2.0	298	2	T27209	hypothetical prote
43	7	2.0	300	1	JC5299	protein kinase (EC
44	7	2.0	300	2	AE0762	conserved hypothet
45	7	2.0	300	2	C84853	hypothetical prote

ALIGNMENTS

RESULT 1
T15523
hypothetical protein C16B8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
C;Accession: T15523
R;Bentley, D.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C16B8.
A;Reference number: Z18365
A;Accession: T15523
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-548 <BEN>
A;Cross-references: EMBL:U41031; NID:g1098982; PID:g1098983; PIDN:AAA82618.1; CESP:C16
C;Genetics:
A;Gene: CESP:C16B8.1
A;Introns: 13/1; 53/3; 123/1; 176/2; 200/1; 229/3; 255/1; 371/1; 466/3; 512/2
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homo

Query Match 2.5%; Score 9; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 AADVWSLGV 249
|||||||||
Db 463 AADVWSLGV 471

RESULT 2
A95405
conserved hypothetical protein SMa2099 [imported] - Sinorhizobium meliloti (strain 102
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: A95405
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bo
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melil
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: A95405
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-190 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65803.1; PID:g14524305; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Huble
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaur
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa2099
A;Genome: plasmid

Query Match 2.2%; Score 8; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 EPAERLTA 306
||||||||
Db 11 EPAERLTA 18

RESULT 3
T11854
protein kinase (EC 2.7.1.-) - Trypanosoma brucei
C;Species: Trypanosoma brucei
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C;Accession: T11854
R;Gale, M.J.; Parsons, M.
Mol. Biochem. Parasitol. 59, 111-122, 1993
A;Title: Trypanosoma brucei gene family encoding protein kinases with catalytic domains
A;Reference number: Z17363; MUID:93295429; PMID:8515773
A;Accession: T11854
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-431 <GAL>
A;Cross-references: EMBL:L03778; NID:g162169; PID:g162170
C;Genetics:
A;Note: nrkA
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP; phosphotransferase

Query Match 2.2%; Score 8; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 242 ADVWSLGV 249
||||||||
Db 205 ADVWSLGV 212

RESULT 4
A83266
conserved hypothetical protein PA3048 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: A83266
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83266
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-725 <STO>
A;Cross-references: GB:AE004729; GB:AE004091; NID:g9949143; PIDN:AAG06436.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA3048

Query Match 2.2%; Score 8; DB 2; Length 725;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 339 GLGLDEAR 346
||||||||
Db 25 GLGLDEAR 32

RESULT 5
AH3258
cation-transporting ATPase pacs (EC 3.6.1.-) [imported] - Brucella melitensis (strain 1
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AH3258
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lete
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melite
A;Reference number: AD3252; PMID:11756688
A;Accession: AH3258
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-813 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL51235.1; PID:g17981923; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0053
A;Map position: I
C;Superfamily: Bacillus probable copper-transporting ATPase yvgX; ATPase nucleotide-bi
C;Keywords: hydrolase

Query Match 2.2%; Score 8; DB 2; Length 813;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 AAPAGSLS 14
||||||||
Db 188 AAPAGSLS 195

RESULT 6
T36211
probable excisionase - Streptomyces coelicolor (fragment)
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C;Accession: T36211
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21577
A;Accession: T36211
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-33 <OLI>
A;Cross-references: EMBL:AL049573; PIDN:CAB40309.1; GSPDB:GN00070; SCOEDB:SCE39.01c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCE39.01c
C;Superfamily: excisionase

Query Match 2.0%; Score 7; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 ATAVATA 65
|||||||
Db 4 ATAVATA 10

RESULT 7
H44057
early E3 10.3K protein - human adenovirus 11 (strain BC34) (fragment)
C;Species: Mastadenovirus h11 (human adenovirus 11)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 12-Apr-1995
C;Accession: H44057
R;Mei, Y.F.; Wadell, G.
Virology 191, 125-133, 1992
A;Title: The nucleotide sequence of adenovirus type 11 early 3 region: comparison of g
A;Reference number: A44057; MUID:93033102; PMID:1413499
A;Accession: H44057

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-66 <MEI>
A;Cross-references: GB:M94459
C;Superfamily: adenovirus early E3 10.3K protein

Query Match 2.0%; Score 7; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 TAVATAS 66
||||||
Db 22 TAVATAS 28

RESULT 8
T36199
probable excisionase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000
C;Accession: T36199
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21600
A;Accession: T36199
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-66 <SAU>
A;Cross-references: EMBL:AL035707; PIDN:CAB38895.1; GSPDB:GN00070; SCOEDB:SCE29.20c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCE29.20c
C;Superfamily: excisionase

Query Match 2.0%; Score 7; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 ATAVATA 65
|||||||
Db 4 ATAVATA 10

RESULT 9
AD3056
hypothetical protein Atu4065 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AD3056
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AD3056
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-72 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44866.1; PID:g17742513; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4065
A;Map position: linear chromosome

Query Match 2.0%; Score 7; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 278 YALPAGL 284
|||||||
Db 25 YALPAGL 31

RESULT 10
H89990
hypothetical protein SA1809 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: H89990
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Og
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: H89990
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-80 <KUR>
A;Cross-references: GB:BA000018; PID:g13701796; PIDN:BAB43089.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA1809

Query Match 2.0%; Score 7; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 196 KKLVLEN 202
|||||||
Db 10 KKLVLEN 16

RESULT 11
S78725
protein YKL053c-a - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
A;Variety: strain S288C
C;Date: 15-Jan-1999 #sequence_revision 15-Jan-1999 #text_change 19-Apr-2002
C;Accession: S78725; S78074
R;Rasmussen, S.; von Wettstein, D.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37872
A;Accession: S78725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-86 <RAS>
A;Cross-references: EMBL:Z28054; MIPS:YKL053c-a
A;Accession: S78074
A;Molecule type: DNA
A;Residues: 1-86 <RA2>
A;Cross-references: EMBL:Z28052; NID:g486071; PID:e1283516; PID:g2980812; MIPS:YKL053c
C;Genetics:
A;Cross-references: SGD:S0007243
A;Map position: 11L
A;Note: YKL053c-a

Query Match 2.0%; Score 7; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 342 LDEAREE 348
|||||||
Db 66 LDEAREE 72

RESULT 12
ERAD27
early E3 10.3K protein - human adenovirus 3
C;Species: Mastadenovirus h3 (human adenovirus 3)
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 07-Nov-1997
C;Accession: C31830; H29500

R;Signaes, C.; Akusjaervi, G.; Pettersson, U.
Gene 50, 173-184, 1986
A;Title: Region E3 of human adenoviruses; differences between the oncogenic adenovirus-3
A;Reference number: A91566; MUID:87219876; PMID:3582978
A;Accession: C31830
A;Molecule type: DNA
A;Residues: 1-91 <SIG>
A;Cross-references: GB:M15952; NID:g209901
A;Note: the GenBank entry ADRE3AA PID:g209908 differs from the published sequence in hav
C;Superfamily: adenovirus early E3 10.3K protein
C;Keywords: early protein; transmembrane protein

Query Match 2.0%; Score 7; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 TAVATAS 66
|||||||
Db 22 TAVATAS 28

RESULT 13
ERAD77
early E3 10.3K protein - human adenovirus 7 (strain Gomen)
C;Species: Mastadenovirus h7 (human adenovirus 7)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C;Accession: G31830; S52804; C31830; H29500
R;Hong, J.S.; Mullis, K.G.; Engler, J.A.
Virology 167, 545-553, 1988
A;Title: Characterization of the early region 3 and fiber genes of Ad7.
A;Reference number: A94386; MUID:89073758; PMID:2849239
A;Accession: G31830
A;Molecule type: DNA
A;Residues: 1-91 <HON>
A;Cross-references: GB:M23696; NID:g341012; PIDN:AAA53251.1; PID:g576458
R;Kajon, A.E.; Wadell, G.
submitted to the EMBL Data Library, April 1995
A;Description: Sequence analysis of the E3 region and fiber gene of human adenovirus 7h.
A;Reference number: S52798
A;Accession: S52804
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-91 <KAJ>
A;Cross-references: EMBL:Z48954; NID:g762955; PIDN:CAA88813.1; PID:g762962
C;Superfamily: adenovirus early E3 10.3K protein
C;Keywords: early protein; transmembrane protein

Query Match 2.0%; Score 7; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 TAVATAS 66
|||||||
Db 22 TAVATAS 28

RESULT 14
QQPSHT
hypothetical protein merT - Pseudomonas aeruginosa transposon Tn501
C;Species: Pseudomonas aeruginosa
C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 24-Sep-1999
C;Accession: A04457
R;Misra, T.K.; Brown, N.L.; Fritzinger, D.C.; Pridmore, R.D.; Barnes, W.M.; Haberstroh,
Proc. Natl. Acad. Sci. U.S.A. 81, 5975-5979, 1984
A;Title: Mercuric ion-resistance operons of plasmid R100 and transposon Tn501: the begin
A;Reference number: A03556; MUID:85014891; PMID:6091128
A;Accession: A04457
A;Molecule type: DNA
A;Residues: 1-116 <MIS>
A;Cross-references: GB:Z00027; GB:K00031; GB:K01725; GB:X01297; GB:X03406; NID:g43714; P
C;Genetics:
A;Gene: merT
C;Superfamily: merT protein
C;Keywords: transmembrane protein

Query Match 2.0%; Score 7; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 LAVLEPY 111
|||||||
Db 44 LAVLEPY 50

RESULT 15
B69418
hypothetical protein AF1347 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: B69418
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dods
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archa
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: B69418
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-123 <KLE>
A;Cross-references: GB:AE001011; GB:AE000782; NID:g2689334; PIDN:AAB89916.1; PID:g2649

Query Match 2.0%; Score 7; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 VLLEPEE 78
|||||||
Db 5 VLLEPEE 11

Search completed: January 15, 2004, 15:05:54
Job time : 22 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:05:58 ; Search time 37 Seconds
(without alignments)
1978.494 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 358
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 762491 seqs, 204481190 residues

Word size : 0

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	274	76.5	358	9	US-09-799-875-8	Sequence 8, Appli
2	234	65.4	360	15	US-10-024-828-9	Sequence 9, Appli
3	208	58.1	233	9	US-09-925-301-1102	Sequence 1102, Ap
4	62	17.3	153	9	US-09-925-301-1367	Sequence 1367, Ap
5	14	3.9	138	9	US-09-864-761-45767	Sequence 45767, A
6	13	3.6	206	15	US-10-228-263-2	Sequence 2, Appli
7	13	3.6	269	12	US-10-291-172-304	Sequence 304, App
8	9	2.5	97	9	US-09-864-761-45703	Sequence 45703, A
9	9	2.5	290	12	US-10-291-172-680	Sequence 680, App
10	9	2.5	548	12	US-10-369-493-6754	Sequence 6754, Ap
11	8	2.2	159	15	US-10-156-761-12825	Sequence 12825, A
12	8	2.2	357	15	US-10-156-761-13581	Sequence 13581, A
13	8	2.2	725	9	US-09-815-242-5125	Sequence 5125, Ap
14	8	2.2	1352	15	US-10-156-761-9867	Sequence 9867, Ap
15	8	2.2	1701	10	US-09-963-959-2	Sequence 2, Appli
16	7	2.0	8	12	US-10-348-232-168	Sequence 168, App
17	7	2.0	31	12	US-10-029-386-32285	Sequence 32285, A
18	7	2.0	35	9	US-09-864-761-46811	Sequence 46811, A
19	7	2.0	61	12	US-10-029-386-31085	Sequence 31085, A
20	7	2.0	90	15	US-10-083-357-740	Sequence 740, App
21	7	2.0	102	11	US-09-764-891-3529	Sequence 3529, Ap
22	7	2.0	102	11	US-09-764-891-3951	Sequence 3951, Ap
23	7	2.0	114	12	US-10-264-049-3683	Sequence 3683, Ap
24	7	2.0	117	9	US-09-799-875-30	Sequence 30, Appl
25	7	2.0	130	9	US-09-799-875-22	Sequence 22, Appl
26	7	2.0	158	12	US-10-104-047-3610	Sequence 3610, Ap
27	7	2.0	158	12	US-10-108-260A-4566	Sequence 4566, Ap
28	7	2.0	160	11	US-09-984-271-211	Sequence 211, App
29	7	2.0	170	12	US-10-104-047-2785	Sequence 2785, Ap
30	7	2.0	176	9	US-09-799-875-27	Sequence 27, Appl
31	7	2.0	185	15	US-10-156-761-8607	Sequence 8607, Ap
32	7	2.0	186	9	US-09-910-150-14	Sequence 14, Appl
33	7	2.0	201	10	US-09-901-540-3	Sequence 3, Appli
34	7	2.0	201	10	US-09-896-856-8	Sequence 8, Appli
35	7	2.0	201	10	US-09-901-257-3	Sequence 3, Appli
36	7	2.0	201	12	US-10-407-303-8	Sequence 8, Appli
37	7	2.0	201	15	US-10-107-931-8	Sequence 8, Appli
38	7	2.0	201	15	US-10-212-793-4	Sequence 4, Appli
39	7	2.0	212	9	US-09-834-496A-4	Sequence 4, Appli
40	7	2.0	214	10	US-09-515-806-10	Sequence 10, Appl
41	7	2.0	214	10	US-09-860-352A-6	Sequence 6, Appli
42	7	2.0	214	12	US-10-369-493-7805	Sequence 7805, Ap
43	7	2.0	214	12	US-10-410-764-93	Sequence 93, Appl
44	7	2.0	232	11	US-09-809-391-623	Sequence 623, App
45	7	2.0	232	12	US-09-882-171-623	Sequence 623, App

ALIGNMENTS

RESULT 1
US-09-799-875-8
; Sequence 8, Application US/09799875
; Patent No. US20020034780A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. US20020034780A1el Human Protein Kinases and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-875-8

Query Match 76.5%; Score 274; DB 9; Length 358;
Best Local Similarity 100.0%; Pred. No. 1.5e-247;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDM 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDM 144

Qy         145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 204

Qy         205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 264
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 264

Qy        265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAP 324
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAP 324

Qy        325 TRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
                ||||||||||||||||||||||||||||||||||
Db        325 TRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358


RESULT 2
US-10-024-828-9
; Sequence 9, Application US/10024828
; Publication No. US20030036051A1
; GENERAL INFORMATION:
;  APPLICANT: Virca, Duke
;  APPLICANT:  Bird, Timothy A.
;  APPLICANT:  Anderson, Dirk M.
;  APPLICANT:  Marken, John S.
;  TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase
;  TITLE OF INVENTION:  Functions
;  FILE REFERENCE: 2877-US
;  CURRENT APPLICATION NUMBER: US/10/024,828
;  CURRENT FILING DATE:  2001-12-18
;  PRIOR APPLICATION NUMBER: US/09/509,902A
;  PRIOR FILING DATE: 1999-08-03
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 360
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-024-828-9

  Query Match             65.4%;  Score 234;  DB 15;  Length 360;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-210;
  Matches  234;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDM 144
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDM 186

Qy        145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 204
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 246

Qy        205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 264
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 306

Qy        265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD 318
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD 360


RESULT 3
US-09-925-301-1102
; Sequence 1102, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1102
;   LENGTH: 233
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-301-1102

  Query Match             58.1%;  Score 208;  DB 9;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-186;
  Matches  208;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        151 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 210
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 85

Qy        211 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 270
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 145

Qy        271 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 330
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 205

Qy        331 EAAQVVPDGLGLDEAREEEGDREVVLYG 358
                ||||||||||||||||||||||||||||
Db        206 EAAQVVPDGLGLDEAREEEGDREVVLYG 233


RESULT 4
US-09-925-301-1367
; Sequence 1367, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1367
;   LENGTH: 153
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (136)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (138)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (141)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (142)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (143)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (152)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1367

  Query Match             17.3%;  Score 62;  DB 9;  Length 153;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-50;
  Matches   62;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          74 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 133

Qy          61 AV 62
               ||
Db         134 AV 135


RESULT 5
US-09-864-761-45767
; Sequence 45767, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45767
;   LENGTH: 138
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC009486.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.93
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
;   OTHER INFORMATION: SWISSPROT HIT: O74536, EVALUE 1.00e-18
;   OTHER INFORMATION: EST_HUMAN HIT: BE897149.1, EVALUE 2.00e-71
US-09-864-761-45767

  Query Match             3.9%;  Score 14;  DB 9;  Length 138;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        236 SYSGKAADVWSLGV 249
              ||||||||||||||
Db         41 SYSGKAADVWSLGV 54


RESULT 6
US-10-228-263-2
; Sequence 2, Application US/10228263
; Publication No. US20030099985A1
; GENERAL INFORMATION:
;  APPLICANT: LI, Jing
;  TITLE OF INVENTION: AMPLIFIED GENE INVOLVED IN CANCER
;  FILE REFERENCE: 38002-0034
;  CURRENT APPLICATION NUMBER: US/10/228,263
;  CURRENT FILING DATE:  2002-12-04
;  PRIOR APPLICATION NUMBER: US 60/330,797
;  PRIOR FILING DATE: 2001-10-31
;  PRIOR APPLICATION NUMBER: US 60/314,655
;  PRIOR FILING DATE: 2001-08-27
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 206
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-228-263-2

  Query Match             3.6%;  Score 13;  DB 15;  Length 206;
  Best Local Similarity   100.0%;  Pred. No. 0.00075;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVWSLGV 249
              |||||||||||||
Db         94 YSGKAADVWSLGV 106


RESULT 7
US-10-291-172-304
; Sequence 304, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-045
;  CURRENT APPLICATION NUMBER: US/10/291,172
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 09/693,267
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/665,363
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: 09/616,847
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 09/596,193
;  PRIOR FILING DATE: 2000-06-17
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 304
;   LENGTH: 269
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-291-172-304

  Query Match             3.6%;  Score 13;  DB 12;  Length 269;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.00094;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVWSLGV 249
              |||||||||||||
Db        157 YSGKAADVWSLGV 169


RESULT 8
US-09-864-761-45703
; Sequence 45703, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45703
;   LENGTH: 97
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC009486.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.62
;   OTHER INFORMATION: SWISSPROT HIT: P57058, EVALUE 1.00e-06
;   OTHER INFORMATION: EST_HUMAN HIT: AU127403.1, EVALUE 4.00e-51
US-09-864-761-45703

  Query Match             2.5%;  Score 9;  DB 9;  Length 97;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        177 GLVLRDLKL 185
              |||||||||
Db         80 GLVLRDLKL 88


RESULT 9
US-10-291-172-680
; Sequence 680, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: No. US20030228584Alel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-045
;  CURRENT APPLICATION NUMBER: US/10/291,172
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 09/693,267
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/665,363
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: 09/616,847
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 09/596,193
;  PRIOR FILING DATE: 2000-06-17
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 680
;   LENGTH: 290
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-291-172-680

  Query Match             2.5%;  Score 9;  DB 12;  Length 290;
  Best Local Similarity   100.0%;  Pred. No. 5.5;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        237 YSGKAADVW 245
              |||||||||
Db        168 YSGKAADVW 176


RESULT 10
US-10-369-493-6754
; Sequence 6754, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6754
;   LENGTH: 548
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-10-369-493-6754
```

Query Match 2.5%; Score 9; DB 12; Length 548;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        241 AADVWSLGV 249
              |||||||||
Db        463 AADVWSLGV 471
```

RESULT 11
US-10-156-761-12825
; Sequence 12825, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12825
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12825

Query Match 2.2%; Score 8; DB 15; Length 159;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        338 DGLGLDEA 345
              ||||||||
Db        125 DGLGLDEA 132
```

RESULT 12
US-10-156-761-13581
; Sequence 13581, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13581
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13581

Query Match 2.2%; Score 8; DB 15; Length 357;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        246 SLGVALFT 253
              ||||||||
Db        108 SLGVALFT 115
```

RESULT 13
US-09-815-242-5125
; Sequence 5125, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5125
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5125

Query Match 2.2%; Score 8; DB 9; Length 725;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        339 GLGLDEAR 346
              ||||||||
Db         25 GLGLDEAR 32
```

RESULT 14
US-10-156-761-9867
; Sequence 9867, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761

```
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9867
;   LENGTH: 1352
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9867

  Query Match             2.2%;  Score 8;  DB 15;  Length 1352;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        301 AERLTATG 308
              ||||||||
Db        846 AERLTATG 853


RESULT 15
US-09-963-959-2
; Sequence 2, Application US/09963959
; Patent No. US20020165145A1
; GENERAL INFORMATION:
;  APPLICANT: Meyers, Rachel A.
;  TITLE OF INVENTION: 33521, A NOVEL HUMAN GUANINE NUCLEOTIDE EXCHANGE FAMILY MEMBER
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: 10448-095001
;  CURRENT APPLICATION NUMBER: US/09/963,959
;  CURRENT FILING DATE:  2001-09-25
;  PRIOR APPLICATION NUMBER: 60/235,033
;  PRIOR FILING DATE: 2000-09-25
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1701
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-963-959-2

  Query Match             2.2%;  Score 8;  DB 10;  Length 1701;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GSLSRKKR 18
              ||||||||
Db        388 GSLSRKKR 395


Search completed: January 15, 2004, 15:11:57
Job time : 38 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 14:58:52 ; Search time 43 Seconds
(without alignments)
1321.491 Million cell updates/sec

Title: US-09-909-474D-2
Perfect score: 358
Sequence: 1 MRATPLAAPAGSLSRKKRLE.....GLGLDEAREEEGDREVVLYG 358

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	274	76.5	358	22	AAU03509	Human protein kina
2	261	72.9	358	22	AAM38908	Human polypeptide
3	261	72.9	393	22	AAM40694	Human polypeptide
4	234	65.4	360	21	AAY69157	PeptideA JJ503-KS
5	208	58.1	233	21	AAB43657	Human cancer assoc
6	208	58.1	358	22	AAB20326	Human protein phos
7	163	45.5	360	23	ABB80976	Human tribbles hom
8	115	32.1	323	22	AAB85791	Human kinase PKIN-
9	83	23.2	278	23	ABB06093	Human NS protein s
10	83	23.2	278	23	ABB06108	Human NS protein s
11	62	17.3	153	21	AAB43922	Human cancer assoc
12	37	10.3	178	22	ABG04413	Novel human diagno
13	14	3.9	138	22	ABG58479	Human liver peptid
14	14	3.9	138	22	ABB43079	Peptide #10585 enc
15	14	3.9	138	22	AAM63978	Human brain expres
16	14	3.9	138	22	AAM76798	Human bone marrow
17	14	3.9	138	22	AAM36904	Peptide #10941 enc
18	14	3.9	138	23	ABG45957	Human peptide enco
19	13	3.6	206	24	ABP96856	Human C8FW protein
20	13	3.6	269	22	AAU28135	Novel human secret
21	13	3.6	372	23	ABB80975	Human tribbles hom
22	9	2.5	97	22	ABG58378	Human liver peptid
23	9	2.5	97	22	ABB42968	Peptide #10474 enc
24	9	2.5	97	22	AAM63872	Human brain expres
25	9	2.5	97	22	AAM76686	Human bone marrow
26	9	2.5	97	22	AAM36793	Peptide #10830 enc
27	9	2.5	97	23	ABG45892	Human peptide enco
28	9	2.5	290	22	AAU28323	Novel human secret
29	8	2.2	725	22	AAU33629	Pseudomonas aerugi
30	8	2.2	1159	22	ABB63460	Drosophila melanog
31	8	2.2	1701	23	ABB08024	Human Rho GEF memb
32	8	2.2	1725	23	ABG91809	Human intracellula
33	7	2.0	35	22	ABG52461	Human liver peptid
34	7	2.0	35	22	ABB37647	Peptide #5153 enco
35	7	2.0	35	22	AAM18595	Peptide #5029 enco
36	7	2.0	51	22	AAU54213	Propionibacterium
37	7	2.0	65	22	AAM88522	Human immune/haema
38	7	2.0	65	23	ABP32052	Human ORF1025 prot
39	7	2.0	67	22	AAM80256	Human protein SEQ
40	7	2.0	73	22	AAM89922	Human immune/haema
41	7	2.0	79	22	AAU45811	Propionibacterium
42	7	2.0	87	23	ABP05531	Human ORFX protein
43	7	2.0	102	22	ABB95990	Human testicular a
44	7	2.0	102	22	AAM94871	Human reproductive
45	7	2.0	102	22	AAM95293	Human reproductive

ALIGNMENTS

RESULT 1
AAU03509
ID AAU03509 standard; Protein; 358 AA.
XX
AC AAU03509;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human protein kinase #9.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Planagan P, Clary D;
XX
DR WPI; 2001-343950/36.
DR N-PSDB; AAS06709.

```
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections -
XX
PS   Claim 7; Figure 2; 433pp; English.
XX
CC   AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC   novel protein kinases have been identified as members of the tyrosine
CC   or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be
CC   used for gene therapy and as DNA probes in diagnostic assays.
CC   The protein kinase polypeptides may be used as antigens in the production
CC   of antibodies against the protein kinases and in assays to identify
CC   modulators of protein kinase expression and activity.
XX
SQ   Sequence   358 AA;

  Query Match             76.5%;  Score 274;  DB 22;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 4e-258;
  Matches  274;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDM 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDM 144

Qy        145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 204

Qy        205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 264
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDS 264

Qy        265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAP 324
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAP 324

Qy        325 TRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
               ||||||||||||||||||||||||||||||||||
Db        325 TRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358


RESULT 2
AAM38908
ID   AAM38908 standard; Protein; 358 AA.
XX
AC   AAM38908;.
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2053.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58064.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 3; SEQ ID NO 2053; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   358 AA;

  Query Match             72.9%;  Score 261;  DB 22;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-245;
  Matches  261;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         98 VYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEP 157
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 VYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEP 157

Qy        158 EAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLW 217
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 EAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLW 217

Qy        218 DKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGA 277
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 DKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGA 277

Qy        278 YALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVP 337
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 YALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVP 337

Qy        338 DGLGLDEAREEEGDREVVLYG 358
               |||||||||||||||||||||
Db        338 DGLGLDEAREEEGDREVVLYG 358


RESULT 3
AAM40694
```

```
ID   AAM40694 standard; Protein; 393 AA.
XX
AC   AAM40694;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 5625.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI59850.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 2; SEQ ID NO 5625; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   393 AA;

  Query Match             72.9%;  Score 261;  DB 22;  Length 393;
  Best Local Similarity   100.0%;  Pred. No. 2e-245;
  Matches  261;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         98 VYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEP 157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 VYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEP 192

Qy        158 EAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLW 217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 EAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLW 252

Qy        218 DKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGA 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 DKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGA 312

Qy        278 YALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVP 337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 YALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVP 372

Qy        338 DGLGLDEAREEEGDREVVLYG 358
              |||||||||||||||||||||
Db        373 DGLGLDEAREEEGDREVVLYG 393
```

RESULT 4

```
AAY69157
ID   AAY69157 standard; Peptide; 360 AA.
XX
AC   AAY69157;
XX
DT   30-MAY-2000  (first entry)
XX
DE   PeptideA JJ503-KS comprising domains VIA to XI of a protein kinase.
XX
KW   Kinase activity; molecular weight marker; isoelectric focusing marker;
KW   peptide fragmentation control; cellular signal transduction.
XX
OS   Homo sapiens.
XX
PN   WO200008180-A2.
XX
PD   17-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US17630.
XX
PR   04-AUG-1998;   98US-0095270.
PR   11-SEP-1998;   98US-0099972.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Virca GD,  Bird TA,  Anderson DM,  Marken JS;
XX
DR   WPI; 2000-195584/17.
DR   N-PSDB; AAZ61155.
XX
PT   New human kinase polypeptides and polynucleotides used as molecular
PT   weight markers and as controls for peptide fragmentation  -
XX
PS   Claim 2; Page 10; 60pp; English.
XX
CC   The present sequence represents a partial polypeptide which has kinase
CC   activity. The kinase polynucleotides can be used to express the
CC   polypeptides, and as probes to identify nucleic acids encoding
CC   proteins having kinase activity. The kinase polypeptides and
CC   fragmented polypeptides are used as molecular weight and isoelectric
CC   focusing markers, and as controls for peptide fragmentation. They also
CC   have a number of therapeutic uses as kinases play a central role in
CC   cellular signal transduction. The polypeptides could also be used to
CC   identify binding partner proteins. The polypeptides can also be used as
CC   a reagent to identify any proteins that the polypeptide regulates, and
CC   proteins with which it might interact. The polypeptides may also be
CC   used for preparation of antibodies. The antibodies can be used in
CC   assays to detect the presence of the protein, and to purify the protein
CC   by immunoaffinity chromatography.
XX
SQ   Sequence   360 AA;

  Query Match             65.4%;  Score 234;  DB 21;  Length 360;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-219;
  Matches  234;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy         85 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDM 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 ALHCPTGTEYTCKVYPVQEALAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDM 186

Qy        145 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 HSLVRSRHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLE 246

Qy        205 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPPQDS 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 DSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPPQDS 306

Qy        265 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 EPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQD 360
```

RESULT 5
AAB43657
ID AAB43657 standard; Protein; 233 AA.
XX
AC AAB43657;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1102.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR N-PSDB; AAC77866.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 11; Page 1707-1708; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerary; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 233 AA;

Query Match 58.1%; Score 208; DB 21; Length 233;
Best Local Similarity 100.0%; Pred. No. 5.2e-194;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        151 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 85

Qy        211 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPPQDSEPVLLF 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPPQDSEPVLLF 145

Qy        271 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 205

Qy        331 EAAQVVPDGLGLDEAREEEGDREVVLYG 358
              ||||||||||||||||||||||||||||
Db        206 EAAQVVPDGLGLDEAREEEGDREVVLYG 233
```

RESULT 6
AAB20326
ID AAB20326 standard; Protein; 358 AA.
XX
AC AAB20326;
XX
DT 29-MAY-2001 (first entry)
XX
DE Human protein phosphatase and kinase protein-5.
XX
KW Protein phosphatase and kinase protein; PPHKP-5; human;
KW gastrointestinal disorder; immune system disorder;
KW neurological disorder; cell proliferative disorder; cancer;
KW diagnosis; therapy.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Region	142..315
FT		/note= "protein kinase signature"
FT	Domain	242..264
FT		/note= "tyrosine kinase catalytic domain signature"
FT	Domain	78..303
FT		/note= "protein kinase domain signature"
FT	Modified-site	54
FT		/note= "O-phosphorylated"
FT	Modified-site	90
FT		/note= "O-phosphorylated"
FT	Modified-site	140
FT		/note= "O-phosphorylated"
FT	Modified-site	210
FT		/note= "O-phosphorylated"
FT	Modified-site	215
FT		/note= "O-phosphorylated"
FT	Modified-site	14
FT		/note= "O-phosphorylated"
FT	Modified-site	27
FT		/note= "O-phosphorylated"
FT	Modified-site	95
FT		/note= "O-phosphorylated"

```
FT   Modified-site   232
FT                   /note= "O-phosphorylated"
FT   Modified-site   238
FT                   /note= "O-phosphorylated"
XX
PN   WO200120004-A2.
XX
PD   22-MAR-2001.
XX
PP   14-SEP-2000; 2000WO-US25515.
XX
PR   15-SEP-1999;   99US-0154141.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Tang YT,  Bandman O,  Hillman JL,  Baughn MR,  Azimzai Y;
PI   Lu DAM;
XX
DR   WPI; 2001-244811/25.
DR   N-PSDB; AAF30480.
XX
PT   Novel human protein phosphatase and kinase proteins for diagnosis,
PT   treatment and prevention of gastrointestinal, immune system,
PT   neurological and cell proliferative disorders -
XX
PS   Claim 1; Page 88-89; 103pp; English.
XX
CC   The present sequence is that of novel human protein phosphatase
CC   and kinase protein PPHKP-5, as predicted from Incyte Clone ID No.
CC   1271505CB1 (see AAF30480). Tissues that express PPHKP-5 (as a
CC   fraction of total tissues expressing PPHKP-5) include reproductive
CC   (0.288), gastrointestinal (0.212) and haematopoietic or immune
CC   (0.192). Diseases or conditions associated with tissues expressing
CC   PPHKP-5 (as a fraction of total tissues expressing PPHKP-5) include
CC   cancer (0.577), inflammation or trauma (0.327) and cell
CC   proliferation (0.308). The encoded protein shows homology to rat
CC   kinase. The invention provides human PPHKP-1 to -11 polypeptides
CC   (see AAB20322-32) and polynucleotides (see AAF30476-86). It also
CC   provides expression vectors, host cells, antibodies, agonists and
CC   antagonists, as well as methods for diagnosing, treating or
CC   preventing disorders associated with expression of PPHKP, including
CC   gastrointestinal disorders, immune system disorders, neurological
CC   disorders and cell proliferative disorders, including cancer.
XX
SQ   Sequence   358 AA;

  Query Match             58.1%;  Score 208;  DB 22;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-194;
  Matches  208;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        151 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 RHRIPEPEAAVLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLT 210

Qy        211 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLF 270

Qy        271 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 GKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLW 330

Qy        331 EAAQVVPDGLGLDEAREEEGDREVVLYG 358
              ||||||||||||||||||||||||||||
Db        331 EAAQVVPDGLGLDEAREEEGDREVVLYG 358


RESULT 7
ABB80976
ID   ABB80976 standard; Protein; 360 AA.
XX
AC   ABB80976;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human tribbles homologue-3 (htrb-3) polypeptide encoding DNA.
XX
KW   Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic; osteopathic;
KW   opthalamological; cardiant; cytostatic; haemostatic; immunosuppressive;
KW   antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
KW   tumour necrosis factor; TNF; htrb-3.
XX
OS   Homo sapiens.
XX
PN   WO200253743-A2.
XX
PD   11-JUL-2002.
XX
PF   08-JAN-2002; 2002WO-US00070.
XX
PR   08-JAN-2001; 2001US-260294P.
XX
PA   (INTE-) INTERLEUKIN GENETICS INC.
XX
PI   Dower S,  Quanstrom E,  Kiss-Toth E;
XX
DR   WPI; 2002-590635/63.
DR   N-PSDB; ABN86479.
XX
PT   Novel isolated human tribbles homologue-1 polypeptide for inhibiting
PT   AP-1-mediated inflammatory signal in a cell, and activating
PT   ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
PT   -
XX
PS   Example 10; Fig 11B; 131pp; English.
XX
CC   The invention provides an isolated human tribbles homologue-1 (htrb-1,
CC   also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC   htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC   signal in a cell. The polypeptide employed in the method is preferably
CC   htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or
CC   htrb-3 N C. It is also useful for providing htrb agonist activity for
CC   activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC   signal, an estrogen receptor-mediated gene activation signal, an
CC   fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
CC   in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC   inflammatory signal in a cell such as tumor necrosis factor (TNF)
CC   induced inflammatory signal, or an interleukin induced inflammatory
CC   signal. htrb proteins are useful in screening assays, predictive medicine
CC   and in therapeutics or prophylactics. The htrb proteins are useful for
CC   screening compounds e.g. for treating and/or preventing diseases caused
CC   by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
CC   psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
CC   and cancers. The htrb therapeutics are useful for antagonizing
CC   interleukin-1 dependent disorders of human placenta, intraventricular
CC   hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
CC   and inflammation or autoimmune disorders. The present sequence represents
CC   the htrb-3 polypeptide.
XX
SQ   Sequence   360 AA;

  Query Match             45.5%;  Score 163;  DB 23;  Length 360;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-150;
  Matches  163;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        196 KKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTML 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 KKLVLENLEDSCVLTGPDDSLWDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTML 257

Qy        256 AGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWL 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWL 317

Qy        316 RQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              |||||||||||||||||||||||||||||||||||||||||||
```

```
Db        318 RQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 360

RESULT 8
AAB85791
ID   AAB85791 standard; Protein; 323 AA.
XX
AC   AAB85791;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Human kinase PKIN-10.
XX
KW   PKIN; kinase; cytostatic; immunosuppressive; immunostimulant; human;
KW   antiarteriosclerotic; cardiant; gene therapy; antisense therapy.
XX
OS   Homo sapiens.
XX
PN   WO200160991-A2.
XX
PD   23-AUG-2001.
XX
PF   16-FEB-2001; 2001WO-US05240.
XX
PR   17-FEB-2000; 2000US-0183682.
PR   02-MAR-2000; 2000US-0186559.
PR   09-MAR-2000; 2000US-0188606.
PR   17-MAR-2000; 2000US-0189998.
PR   30-MAR-2000; 2000US-0193851.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Buford N,  Gandhi AR,  Patterson C,  Khan FA,  Yue H;
PI   Hafalia A,  Shih LL,  Tribouley CM,  Yao MG,  Burrill JD,  Marcus GA;
PI   Zingler KA,  Lu DAM,  Bandman O,  Policky JL,  Griffin JA,  Thornton M;
PI   Nguyen DB,  Lal P,  Walsh RT;
XX
DR   WPI; 2001-514771/56.
DR   N-PSDB; AAH76218.
XX
PT   isolated human kinase polypeptides useful in the diagnosis, treatment
PT   and prevention of cancer, immune disorders and disorders affecting
PT   growth and development -
XX
PS   Claim 1; Page 115; 126pp; English.
XX
CC   The invention provides human kinases (PKIN) and polynucleotides encoding
CC   PKIN. The PKIN polypeptides can be expressed using standard recombinant
CC   methodology. The PKIN polypeptides, polynucleotides, modulators and
CC   specific antibodies are useful in the diagnosis, treatment and prevention
CC   of cancer, immune disorders, disorders affecting growth and development,
CC   atherosclerosis, and other cardiovascular diseases, and lipid disorders
CC   and in the assessment of the effects of exogenous compounds on the
CC   expression of nucleic acid sequences of human kinases. The present
CC   sequence represents a human PKIN-10 polypeptide.
XX
SQ   Sequence   323 AA;

  Query Match             32.1%;  Score 115;  DB 22;  Length 323;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-103;
  Matches  115;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        244 VWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAER 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 VWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVRCLLRREPAER 268

Qy        304 LTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 LTATGILLHPWLRQDPMPLAPTRSHLWEAAQVVPDGLGLDEAREEEGDREVVLYG 323


RESULT 9
ABB06093
ID   ABB06093 standard; Protein; 278 AA.
XX
AC   ABB06093;
XX
DT   10-MAY-2002  (first entry)
XX
DE   Human NS protein sequence SEQ ID NO:185.
XX
KW   Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW   antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW   vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW   anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW   antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW   anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW   gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW   contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW   endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW   rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW   inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW   infertility; cardiovascular disease; coagulation disease; hypertension;
KW   ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW   diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW   gastric ulcer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO200206315-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-IL00653.
XX
PR   18-JUL-2000; 2000IL-0137345.
PR   15-DEC-2000; 2000IL-0140354.
XX
PA   (COMP-) COMPUGEN LTD.
XX
PI   Mintz L,  Freilich S,  Bernstein J;
XX
DR   WPI; 2002-155037/20.
DR   N-PSDB; ABL39747.
XX
PT   One hundred and twenty eight novel nucleic acid sequences, useful for
PT   treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS   Claim 6; Page 213-214; 290pp; English.
XX
CC   ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC   encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC   (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC   antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC   vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC   anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC   anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC   immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC   antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC   nootropic and contraceptive activities. The NS can be used in vaccines,
CC   gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC   antibodies from the present invention can be used for treating and
CC   diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC   diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC   cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC   glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC   disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC   disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC   depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC   Alzheimer's disease and as a contraceptive.
XX
SQ   Sequence   278 AA;

  Query Match             23.2%;  Score 83;  DB 23;  Length 278;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-72;
  Matches   83;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 88

Qy         61 AVATASRLGPYVLLEPEEGGRAY 83
              |||||||||||||||||||||||
Db         89 AVATASRLGPYVLLEPEEGGRAY 111
```

RESULT 10
ABB06108
ID ABB06108 standard; Protein; 278 AA.
XX
AC ABB06108;
XX
DT 10-MAY-2002 (first entry)
XX
DE Human NS protein sequence SEQ ID NO:200.
XX
KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO200206315-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-IL00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Freilich S, Bernstein J;
XX
DR WPI; 2002-155037/20.
DR N-PSDB; ABL39762.
XX
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 6; Page 231-232; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SQ Sequence 278 AA;

```
  Query Match             23.2%;  Score 83;  DB 23;  Length 278;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-72;
  Matches   83;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 88

Qy         61 AVATASRLGPYVLLEPEEGGRAY 83
              |||||||||||||||||||||||
Db         89 AVATASRLGPYVLLEPEEGGRAY 111
```

RESULT 11
AAB43922
ID AAB43922 standard; Protein; 153 AA.
XX
AC AAB43922;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1367.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR N-PSDB; AAC78131.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 11; Page 2031; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerary; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The

CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 153 AA;

Query Match 17.3%; Score 62; DB 21; Length 153;
Best Local Similarity 100.0%; Pred. No. 5.4e-52;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLPPCLLPLSPPTAPDRAT 133

Qy         61 AV 62
              ||
Db        134 AV 135
```

RESULT 12
ABG04413
ID ABG04413 standard; Protein; 178 AA.
XX
AC ABG04413;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #4404.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS68600.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 34772; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 178 AA;

Query Match 10.3%; Score 37; DB 22; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.3e-27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         28 ERPVQKRARSGPQPRLPPCLLPLSPPTAPDRATAVAT 64
              |||||||||||||||||||||||||||||||||||||
Db        110 ERPVQKRARSGPQPRLPPCLLPLSPPTAPDRATAVAT 146
```

RESULT 13
ABG58479
ID ABG58479 standard; Peptide; 138 AA.
XX
AC ABG58479;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID No 37127.
XX
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00664.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX
PS Claim 27; SEQ ID No 37127; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (I) may be used for predicting, measuring and displaying gene

```
CC   expression in samples derived from human adult liver. The genes
CC   identified may be involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC   is associated with coronary heart disease. ABG47348-ABG59930 represent
CC   human liver single exon encoded peptides of the invention.
CC   Note: The sequence information for this patent does not appear in the
CC   printed specification but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   138 AA;

  Query Match             3.9%;  Score 14;  DB 22;  Length 138;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        236 SYSGKAADVWSLGV 249
              ||||||||||||||
Db         41 SYSGKAADVWSLGV 54


RESULT 14
ABB43079
ID   ABB43079 standard; Peptide; 138 AA.
XX
AC   ABB43079;
XX
DT   04-FEB-2002  (first entry)
XX
DE   Peptide #10585 encoded by human foetal liver single exon probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 27; SEQ ID NO 35714; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a peptide encoded by a single exon
CC   nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   138 AA;

  Query Match             3.9%;  Score 14;  DB 22;  Length 138;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        236 SYSGKAADVWSLGV 249
              ||||||||||||||
Db         41 SYSGKAADVWSLGV 54


RESULT 15
AAM63978
ID   AAM63978 standard; Protein; 138 AA.
XX
AC   AAM63978;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe encoded protein SEQ ID NO: 36083.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 36083; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is a protein encoded by one of
CC   the probes of the invention.
XX
SQ   Sequence   138 AA;

  Query Match             3.9%;  Score 14;  DB 22;  Length 138;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        236 SYSGKAADVWSLGV 249
              ||||||||||||||
Db         41 SYSGKAADVWSLGV 54


Search completed: January 15, 2004, 15:04:49
Job time : 44 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 10:03:56 ; Search time 2407.84 Seconds
(without alignments)
10840.821 Million cell updates/sec

Title: US-09-909-474D-1_COPY_49_1122
Perfect score: 1074
Sequence: 1 atgcgagccacccctctggc.....gagaagtggttctgtatggc 1074

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	812.4	75.6	863	10	BG748142	BG748142 602705453
2	802.4	74.7	1036	13	BQ217691	BQ217691 AGENCOURT
3	801.4	74.6	873	13	BQ216198	BQ216198 AGENCOURT
4	800.8	74.6	945	13	BX325345	BX325345 BX325345
5	795.6	74.1	902	13	BX421477	BX421477 BX421477
6	790.2	73.6	929	13	BU500619	BU500619 AGENCOURT
7	770.4	71.7	1201	9	AL556690	AL556690 AL556690
8	768.6	71.6	948	13	BX443303	BX443303 BX443303
9	758.2	70.6	918	13	BQ223497	BQ223497 AGENCOURT
10	756	70.4	914	13	BU500176	BU500176 AGENCOURT
11	744	69.3	820	12	BI256421	BI256421 602974439
12	740	68.9	929	9	AL522632	AL522632 AL522632
13	736.4	68.6	892	14	CA495347	CA495347 AGENCOURT
14	717.6	66.8	1126	13	BX405957	BX405957 BX405957
15	711	66.2	944	13	BX363066	BX363066 BX363066
16	707	65.8	1201	13	BX445734	BX445734 BX445734
17	705.8	65.7	1076	13	BX367265	BX367265 BX367265
18	703.4	65.5	838	10	BG171687	BG171687 602322210
19	701.4	65.3	738	12	BM007742	BM007742 603617190
20	696.6	64.9	851	12	BI909168	BI909168 603069177
21	694.4	64.7	913	13	BQ433764	BQ433764 AGENCOURT
22	692.6	64.5	880	10	BG470123	BG470123 602533821
23	688.4	64.1	851	10	BG331647	BG331647 602432667
24	687.8	64.0	1053	13	BQ224844	BQ224844 AGENCOURT
25	687	64.0	989	10	BE792929	BE792929 601585274
26	686	63.9	686	12	BM764986	BM764986 K-EST0046
27	681.2	63.4	766	12	BI253633	BI253633 602975310
28	680.2	63.3	810	12	BM045379	BM045379 603623380
29	679.8	63.3	748	12	BI085445	BI085445 602869933
30	666.6	62.1	931	10	BG470193	BG470193 602533702
31	665.2	61.9	957	13	BQ949069	BQ949069 AGENCOURT
32	661.4	61.6	910	10	BG397691	BG397691 602438746
33	659.4	61.4	891	9	AL555267	AL555267 AL555267
34	655.6	61.0	716	10	BG422518	BG422518 602447002
35	654.4	60.9	703	12	BI334267	BI334267 602997785
36	649.8	60.5	1922	11	AK089931	AK089931 Mus muscu
37	648	60.3	648	12	BM743773	BM743773 K-EST0017
38	647.8	60.3	947	13	BQ928125	BQ928125 AGENCOURT
39	645	60.1	645	12	BM772759	BM772759 K-EST0056
40	641.2	59.7	878	10	BF309552	BF309552 601892066
41	637.6	59.4	686	10	BG750849	BG750849 602707375
42	637.4	59.3	932	13	BQ683898	BQ683898 AGENCOURT
43	636	59.2	648	12	BM791610	BM791610 K-EST0071
44	630.2	58.7	943	13	BQ931607	BQ931607 AGENCOURT
45	627.2	58.4	946	13	BU158759	BU158759 AGENCOURT

ALIGNMENTS

RESULT 1
BG748142
LOCUS BG748142 863 bp mRNA linear EST 15-MAY-2001
DEFINITION 602705453F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4842324 5', mRNA sequence.
ACCESSION BG748142
VERSION BG748142.1 GI:14058795
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 863)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1676 row: i column: 13
High quality sequence stop: 833.

```
FEATURES             Location/Qualifiers
     source          1. .863
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4842324"
                     /tissue_type="normal pigmented retinal epithelium"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_43"
                     /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
BASE COUNT      135 a    300 c    264 g    164 t
ORIGIN

  Query Match             75.6%;  Score 812.4;  DB 10;  Length 863;
  Best Local Similarity   98.6%;  Pred. No. 4.5e-159;
  Matches  830;  Conservative    0;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 82

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 142

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 202

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 262

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 322

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 382

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 442

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 502

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 562

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 622

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 682

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 742

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        743 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAGTACCCC 802

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840
              |||||||||||||| |    |||||||||| ||||||| || |||||||| ||||||
Db        803 TTCCAGGACTCGGAAGCCTGTCTGCTCTTCGGGCAGATCCG-CGTGGGGCCTAAGCCTTG 861

Qy        841 CC 842
              ||
Db        862 CC 863


RESULT 2
BQ217691
LOCUS       BQ217691                1036 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7570945 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6058546
            5', mRNA sequence.
ACCESSION   BQ217691
VERSION     BQ217691.1  GI:20399091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1036)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13323  row: m  column: 11
            High quality sequence stop: 640.
FEATURES             Location/Qualifiers
     source          1. .1036
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6058546"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
BASE COUNT      173 a    365 c    301 g    196 t      1 others
ORIGIN

  Query Match             74.7%;  Score 802.4;  DB 13;  Length 1036;
  Best Local Similarity   97.9%;  Pred. No. 5.7e-157;
  Matches  866;  Conservative    0;  Mismatches   12;  Indels    7;  Gaps    5;

Qy         46 AAGAAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCT 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGAAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCT 60

Qy        106 CGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCT 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCT 120

Qy        166 CCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAG 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAG 180

Qy        226 CCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACC 285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 CCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACC 240

Qy        286 TGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCC 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCC 300

Qy        346 CCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCC 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCC 360

Qy        406 TTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCT 465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCT 420

Qy        466 GAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAG 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAG 480

Qy        526 CACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG 540

Qy        586 AAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCC 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCC 600

Qy        646 CTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCC 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCC 660

Qy        706 TCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCT- 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        661 TCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTG 720

Qy        765 GGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCC--GC 822
              |||||||||||||||| | |||||||||||||||||||||||| |||| |||| ||
Db        721 GGCCGGCCACTACCCCTTTCCGGACTCGGAGCCTGTCCTGCTCTTCCGCAAAATCCCGGC 780

Qy        823 CGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTG-CCCGCTGTCTGGTTCGCTG 881
              ||||||| |||||||||||||||| |||||||||||||| |||| |||| ||||||||||
Db        781 CGCGGGGGCTACGCCTTGCCTGCAAGCCTCTCGGCCCCTGCCCCGCTGTCTGGTTCGCTG 840

Qy        882 CC-TCCTTCGTCGGGA--GCCAGCTGAACGGCTCACAGCCACAGG 923
              || |||||| || |||   ||  ||||||||||||||||||||
Db        841 CCTTCCTTCNTCCGGAAGCCCGCTTGAACGGCTCACAGCCACAGG 885
```

RESULT 3
BQ216198

```
LOCUS       BQ216198                873 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7574831 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059905
            5', mRNA sequence.
ACCESSION   BQ216198
VERSION     BQ216198.1  GI:20397598
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 873)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13327  row: f  column: 02
            High quality sequence stop: 640.
FEATURES             Location/Qualifiers
     source          1. .873
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6059905"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
BASE COUNT      144 a    305 c    264 g    160 t
ORIGIN
```

Query Match 74.6%; Score 801.4; DB 13; Length 873;
Best Local Similarity 99.1%; Pred. No. 8.8e-157;
Matches 816; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 110

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 170

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 230

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 290

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 350

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 410

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 470

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 530

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 590

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 650

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 710

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 770

Qy        721 GCAGCCGATGTCT-GGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCC 779
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        771 GCAGCCGATGTCTGGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCC 830
```

```
Qy        780 CTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGC 822
              ||||||||||||||||||||||||||||||||| || |  | |
Db        831 CTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGAAAAAAACCC 873


RESULT 4
BX325345
LOCUS       BX325345                945 bp    mRNA    linear   EST 02-MAY-2003
DEFINITION  BX325345 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
            Homo sapiens cDNA clone CS0DL004YF04 5-PRIME, mRNA sequence.
ACCESSION   BX325345
VERSION     BX325345.1  GI:30338449
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 945)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AL004DC02QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AL004DC02QP1.
FEATURES              Location/Qualifiers
     source           1. .945
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DL004YF04"
                      /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                      /cell_line="RAMOS CELL LINE"
                      /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                      25-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      163 a    313 c    282 g    176 t     11 others
ORIGIN

  Query Match             74.6%;  Score 800.8;  DB 13;  Length 945;
  Best Local Similarity   97.7%;  Pred. No. 1.2e-156;
  Matches  821;  Conservative   10;  Mismatches    7;  Indels    2;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 166

Qy         61 TTG-GATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCA 119
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TTGAGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCA 226

Qy        120 GCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAAC 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAAC 286

Qy        180 TGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGG 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 TGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGG 346

Qy        240 GCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCC 299
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GCGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCC 406

Qy        300 CGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGT 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGT 466

Qy        360 GGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGAC 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 GGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGAC 526

Qy        420 CCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGC 479
              |||||||||||||||||||||||||||||| :||||||||||||||||||||||||||||
Db        527 CCATGGGGACATGCACAGCCTGGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGC 585

Qy        480 CGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCT 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCT 645

Qy        540 GCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCT 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 GCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCT 705

Qy        600 GGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCA 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 GGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCA 765

Qy        660 CGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAA 719
              ||||||||||||||||||||||||||||||:||||||||||||||||:||||||||||||
Db        766 CGCGTGCCCAGCCTACGTGGGACCTGAGAWACTCAGCTCACGGGCCTYATACTCGGGCAA 825

Qy        720 GGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCC 779
              :| |||||||||:||||||||||||||||||||||||||||||:||||||||||||||||
Db        826 RGGAGCCGATGTYTGGAGCCTGGGCGTGGCGCTCTTCACCAWGCTGGCCGGCCACTACCC 885

Qy        780 CTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTT 839
              ||||||:|||||||||||||| || |||||:||:|| ||:|||||||||| ||| |||||
Db        886 CTTCCARGACTCGGAGCCTGTTCTTCTCTTYGGSAAAATYCGCCGCGGGGGCTAAGCCTT 945


RESULT 5
BX421477
LOCUS       BX421477                902 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX421477 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG007YG03 5-PRIME, mRNA sequence.
ACCESSION   BX421477
VERSION     BX421477.1  GI:30764090
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 902)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DG007AD02QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DG007AD02QP1.
FEATURES              Location/Qualifiers
     source           1. .902
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DG007YG03"
                      /tissue_type="B CELLS (RAMOS CELL LINE)"
```

```
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
BASE COUNT      157 a    304 c    268 g    169 t      4 others
ORIGIN

  Query Match             74.1%;  Score 795.6;  DB 13;  Length 902;
  Best Local Similarity   99.4%;  Pred. No. 1.4e-155;
  Matches  817;  Conservative    2;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 AKGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 142

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 202

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 262

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 322

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| |||||||||||||||||||||||||||||||||||| |||||||||
Db        323 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 381

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 441

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 501

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||
Db        502 CATGGGGACATGCACAGCCTGGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGCC 560

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 620

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 680

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 740

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 800

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 860

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGC 822
              |||||||||||||||||||||||||||||||||||||||||||
Db        861 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGC 902


RESULT 6
BU500619
```

```
LOCUS       BU500619                929 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_7861060 NIH_MGC_64 Homo sapiens cDNA clone IMAGE:6109605
            5', mRNA sequence.
ACCESSION   BU500619
VERSION     BU500619.1  GI:22801804
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: NIMH/LOG
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2355  row: 1  column: 22
            High quality sequence start: 13
            High quality sequence stop: 731.
FEATURES             Location/Qualifiers
     source          1. .929
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6109605"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_64"
                     /note="Vector: pOTB7a; Site_1: CeuI; Site_2: SceI; This
                     library is a size selection of NIH_MGC_35, from 3.0-4.5
                     kb. Size selection done at the National Institute of
                     Mental Health, NIH. Note: this is a NIH_MGC Library."
BASE COUNT      146 a    328 c    281 g    174 t
ORIGIN

  Query Match             73.6%;  Score 790.2;  DB 13;  Length 929;
  Best Local Similarity   95.2%;  Pred. No. 1.9e-154;
  Matches  871;  Conservative    0;  Mismatches   33;  Indels   11;  Gaps    5;

Qy         49 AAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGA 108
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 AGGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGA 69

Qy        109 AGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCA 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 AGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCA 129

Qy        169 GATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCC 189

Qy        229 GAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGC 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGC 249

Qy        289 AAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 AAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCG 309

Qy        349 CACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTT 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTT 369

Qy        409 TTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 TTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAG 429
```

```
Qy        469 CCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCAC 528
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        430 CCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACTGCCCTGGCGCACTGTCACCAGCAC 489

Qy        529 GGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 GGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAG 549

Qy        589 AAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTG 648
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 AAGCTGGTGCTAGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTG 609

Qy        649 TGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCA 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 TGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCA 669

Qy        709 TACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCC 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 TACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCC 729

Qy        769 GGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGA---TCCGCCGC 825
              |||||||||||||||||||||||||||||||||||||||||||||||| ||    |  ||
Db        730 GGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAGGAATCCGCCGCGG 789

Qy        826 GGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCT-GCCCGCTGTCTGGTTCGCTGCC- 883
              ||| |    |||||||||||||||||||||||||| |||| ||| |||| ||||||||
Db        790 GGGCCTACGCCCTTGCCTGCAGGCCTCTCGGCCCCTGGCCCCCTGGCTGGGTCGCTGCCT 849

Qy        884 ----TCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACA-GGCATCCTCCTGCA-CC 937
                  | | |||    |  |     || |||||||||||||| |||||||||||||| ||
Db        850 TCCTTTCGTCGGGAAGGCCAGCTGGACCGGCTCACAGCCACAGGGCATCCTCCTGCACCC 909

Qy        938 CCTGGCTGCGACAGG 952
              ||||||||||| |||
Db        910 CCTGGCTGCGAAAGG 924
```

```
RESULT 7
AL556690
LOCUS       AL556690                1201 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL556690 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK005YC05 5-PRIME, mRNA sequence.
ACCESSION   AL556690
VERSION     AL556690.2  GI:31278491
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12899595.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DK005AB03QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK005AB03QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK005YC05"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      205 a    366 c    336 g    232 t     62 others
ORIGIN
```

```
  Query Match           71.7%;  Score 770.4;  DB 9;  Length 1201;
  Best Local Similarity   96.0%;  Pred. No. 2.7e-150;
  Matches  793;  Conservative   15;  Mismatches   16;  Indels    2;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 245

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 305

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 365

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 425

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 485

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 545

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 605

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||||| :|||||||||||||||||||||||||||
Db        606 CATGGGGACATGCACAGCCTGGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGCC 664

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 724

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTA 784

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        785 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 844

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        845 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 904

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              |||||||||||||||||:|||||||||||:|:||::|||||||| |||||||||
Db        905 GCAGCCGATGTCTGGAGCCTKGGCGTGGCGCTCTTYAMCAWKCTGGCCGGGCACTACCCC 964

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCG 826
              |||:| |||||:|::|:|:|| :  : |  |  ||    ||: | |
Db        965 TTCMA-GACTCSGRRCYTKTCTKCTYTTCGGMAAATCSGCGSGGGG 1009
```

```
RESULT 8
BX443303
LOCUS       BX443303                948 bp    mRNA    linear   EST 22-MAY-2003
```

```
DEFINITION  BX443303 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG007YF05 5-PRIME, mRNA sequence.
ACCESSION   BX443303
VERSION     BX443303.1  GI:31018564
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DG007CC03QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DG007CC03QP1.
FEATURES             Location/Qualifiers
     source          1. .948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG007YF05"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
BASE COUNT      152 a    326 c    296 g    169 t      5 others
ORIGIN

  Query Match             71.6%;  Score 768.6;  DB 13;  Length 948;
  Best Local Similarity   99.4%;  Pred. No. 6e-150;
  Matches  790;  Conservative    2;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 215

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 275

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 335

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 395

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| |||||||||||||||||||||||||||||||||||| |||||||||
Db        396 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 454

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 514

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 574

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||
Db        575 CATGGGGACATGCACAGCCTGGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGCC 633

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        634 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 693

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 753

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        754 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 813

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        814 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 873

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        874 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 933

Qy        781 TTCCAGGACTCGGAG 795
              |||||:|||||||||
Db        934 TTCCARGACTCGGAG 948


RESULT 9
BQ223497
LOCUS       BQ223497                 918 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7576910 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6058986
            5', mRNA sequence.
ACCESSION   BQ223497
VERSION     BQ223497.1  GI:20404897
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 918)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13324  row: o  column: 19
            High quality sequence stop: 684.
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6058986"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
BASE COUNT      144 a    321 c    285 g    168 t
ORIGIN
```

```
  Query Match             70.6%;  Score 758.2;  DB 13;  Length 918;
  Best Local Similarity   98.5%;  Pred. No. 8.7e-148;
  Matches  807;  Conservative    0;  Mismatches    8;  Indels    4;  Gaps    4;

Qy          1 ATGCGAGCCACCCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 ATGCGAGCCACCCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 159

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 219

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 279

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 339

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 399

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 459

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 519

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 579

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 639

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 699

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 759

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        760 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCCGGCAAG 819

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCG-TGGCGCTCTTCACCATGCT-GGCCGGCCACTACC 778
              |||||||||||||||||||||||||| |||||||||||    |||||| ||||||||||||
Db        820 GCAGCCGATGTCTGGAGCCTGGGCGTTGGCGCTCTTTCACATGCTGGGCCGGCCACTACC 879

Qy        779 CC-TTCCAGGACTC-GGAGCCTGTCCTGCTCTTCGGCAA 815
              || |||||||||| |||||||| |||||||||| ||
Db        880 CCTTTCCAGGACTCGGGAGCCTGGTCTGCTCTTCGGGAA 918


RESULT 10
BU500176
LOCUS       BU500176                914 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_7859610 NIH_MGC_64 Homo sapiens cDNA clone IMAGE:6108987
            5', mRNA sequence.
ACCESSION   BU500176
VERSION     BU500176.1  GI:22800682
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 914)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: NIMH/LOG
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2354  row: c  column: 04
            High quality sequence stop: 682.
FEATURES             Location/Qualifiers
     source          1. .914
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6108987"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_64"
                     /note="Vector: pOTB7a; Site_1: CeuI; Site_2: SceI; This
                     library is a size selection of NIH_MGC_35, from 3.0-4.5
                     kb. Size selection done at the National Institute of
                     Mental Health, NIH. Note: this is a NIH_MGC Library."
BASE COUNT      147 a    323 c    274 g    169 t      1 others
ORIGIN

  Query Match             70.4%;  Score 756;  DB 13;  Length 914;
  Best Local Similarity   96.1%;  Pred. No. 2.5e-147;
  Matches  808;  Conservative    0;  Mismatches   26;  Indels    7;  Gaps    3;

Qy         49 AAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGA 60

Qy        109 AGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCA 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCA 120

Qy        169 GATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCC 180

Qy        229 GAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGC 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGC 240

Qy        289 AAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCG 300

Qy        349 CACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTT 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTT 360

Qy        409 TTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAG 420

Qy        469 CCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCAC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCAC 480

Qy        529 GGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAG 540

Qy        589 AAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTG 648
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGCTGGTGCTAGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTG 600
```

```
Qy        649 TGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCA 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCA 660

Qy        709 TACTCGGGCAAGGCAGCCGATGTCTGG-AGCCTGGGCGTGGCGCTCTTCACCATGCTGGC 767
              |||||||||||||||||||||||||| | ||||||||||||||||||||||||||||||
Db        661 TACTCGGGCAAGGCAGCCGATGTCTGGAACCCTGGGCGTGGCGCTCTTCACCATGCTGGC 720

Qy        768 CGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCG---CCG 824
              ||||||||||||||||||||||||||||||||||||||||||||||||| |   |     |
Db        721 CGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAAAAACNGGCCGC 780

Qy        825 CGGGGCCTACGCCTT---GCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTG 881
              ||||||||||||| |    | |||||||||||| | |||     | | |  || || | |
Db        781 GGGGGCCTACGCCCTTGGCCTTGCAGGCCTCTCCGGCCCCTGGCCCCGCGTGCTTGGGTC 840

Qy        882 C 882
              |
Db        841 C 841
```

```
RESULT 11
BI256421
LOCUS       BI256421                820 bp    mRNA    linear   EST 17-JUL-2001
DEFINITION  602974439F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5113905 5',
            mRNA sequence.
ACCESSION   BI256421
VERSION     BI256421.1  GI:14810812
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 820)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11276  row: m  column: 10
            High quality sequence stop: 820.
FEATURES             Location/Qualifiers
     source          1. .820
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5113905"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_12"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
BASE COUNT      133 a    285 c    245 g    157 t
ORIGIN
```

```
  Query Match             69.3%;  Score 744;  DB 12;  Length 820;
  Best Local Similarity   98.7%;  Pred. No. 7.7e-145;
  Matches  813;  Conservative    0;  Mismatches    5;  Indels    6;  Gaps    6;

Qy         43 AGGAAGAAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGA 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGGAAGAAGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGA 60

Qy        103 GCTCGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACT 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACT 120

Qy        163 GCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTG 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTG 180

Qy        223 GAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTAT 282
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 GAGCCCGAGGAGGGCGGGCGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTAT 240

Qy        283 ACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTG 342
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        241 ACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTG 300

Qy        343 CCCCCG-CACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTA 401
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCCCGTCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTA 360

Qy        402 CGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTAT 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTAT 420

Qy        462 CCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCA 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCA 480

Qy        522 CCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGA 581
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        481 CCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTAGTCTTCGCTGACCGTGA 540

Qy        582 GAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGA 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGA 600

Qy        642 TTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACG 701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACG 660

Qy        702 GGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCG-TGGCGCTCTTCACCA 760
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        661 GGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTTGGCGCTCTTCACCA 720

Qy        761 TGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCC 820
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||| ||||||
Db        721 TGCTGGCCGGCCACTA-CCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCCAGATCC 779

Qy        821 GCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCC 864
              | |||||||||||||| |||||||||| ||||||| |||||||
Db        780 G-CGCGGGGCCTACGCC-TGCCTGCAGG-CTCTCGGACCCTGCC 820
```

```
RESULT 12
AL522632
LOCUS       AL522632                929 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AL522632 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
            cDNA clone CS0DB009YO03 5-PRIME, mRNA sequence.
ACCESSION   AL522632
VERSION     AL522632.2  GI:31040900
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 13, 2001 this sequence version replaced gi:12786125.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
```

```
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DB009AH02QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DB009AH02QP1.
FEATURES              Location/Qualifiers
     source           1. .929
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DB009YO03"
                      /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
                      /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      158 a    313 c    284 g    169 t      5 others
ORIGIN

  Query Match             68.9%;  Score 740;  DB 9;  Length 929;
  Best Local Similarity   99.5%;  Pred. No. 5.4e-144;
  Matches  762;  Conservative    1;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 225

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 285

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 345

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 405

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |:|||||||| ||||||||||||||||||||||||||||||||||||| |||||||||
Db        406 CRGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 464

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 524

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 584

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        585 CATGGGGACATGCACAGCCTGGTGCGAAG-CGCCACCGTATCCCTGAGCCTGAGGCTGCC 643

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 703

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 763

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 823

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 883

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGG 766
              |||||||||||||||||||||||||||||||||||||||||||||
Db        884 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGG 929


RESULT 13
CA495347
LOCUS       CA495347                892 bp    mRNA    linear   EST 14-NOV-2002
DEFINITION  AGENCOURT_10853584 MAPcL Homo sapiens cDNA clone IMAGE:6785470 5',
            mRNA sequence.
ACCESSION   CA495347
VERSION     CA495347.1  GI:24958409
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Kristi A. Egland, Ira Pastan
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14286  row: i  column: 21
            High quality sequence stop: 653.
FEATURES              Location/Qualifiers
     source           1. .892
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:6785470"
                      /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1
                      , LNCaP"
                      /lab_host="EMDH10B"
                      /clone_lib="MAPcL"
                      /note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
                      Subtracted with brain, liver, lung, kidney and muscle.
                      Directionally cloned. Priming method: oligo-dT. Average
                      insert size: 1800 bp. Library amplification:  26,000 fold.
                      Kristi A. Egland, James J. Vincent, Robert Strausberg,
                      Bungkook Lee  & Ira Pastan:  Discovery of new breast
                      cancer genes encoding  membrane and secreted proteins.
                      Manuscript submitted."
BASE COUNT      155 a    309 c    267 g    161 t
ORIGIN

  Query Match             68.6%;  Score 736.4;  DB 14;  Length 892;
  Best Local Similarity   97.5%;  Pred. No. 3e-143;
  Matches  780;  Conservative    0;  Mismatches   16;  Indels    4;  Gaps    3;

Qy        215 TCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCA 274
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          1 TCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAGGCCCTGCACTGCCCTACAGGCA 60

Qy        275 CTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACG 334
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         61 CTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATG 120

Qy        335 CGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGC 394
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGC 180

Qy        395 TCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCC 454
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCC 240

Qy        455 ACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGC 514
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGC 300

Qy        515 ACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTG 574
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTG 360

Qy        575 ACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 634
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 420

Qy        635 CAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCA 694
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCA 480

Qy        695 GCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCT 754
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCT 540

Qy        755 TCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCA 814
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCA 600

Qy        815 AGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGG 874
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGG 660

Qy        875 TTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGC 934
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGC 720

Qy        935 -ACCCCTGGCTGCGACAGGA-CCCGATGCCCTTAGCCCCAACCCGA--TCCCATCTCTGG 990
                |||||||||||||||||| |||||||||||| |||||| ||||   ||     | ||
Db        721 AACCCCTGGCTGCGACAGGACCCCGATGCCCTTAACCCCAAACCGAATCCCATCTCCGGG 780

Qy        991 GAGGCTGCCCAGGTGGTCCC 1010
                ||||||||||  | ||  ||
Db        781 GAGGCTGCCCCAGGGGGTCC 800


RESULT 14
BX405957
LOCUS       BX405957                1126 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX405957 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
            CS0DM009YF15 5-PRIME, mRNA sequence.
ACCESSION   BX405957
VERSION     BX405957.1  GI:30768575
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AM009CC08QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AM009CC08QP1.
FEATURES             Location/Qualifiers
     source          1. .1126
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM009YF15"
                     /tissue_type="FETAL LIVER"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL LIVER"
                     /note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
BASE COUNT      176 a    360 c    350 g    211 t     29 others
ORIGIN

  Query Match             66.8%;  Score 717.6;  DB 13;  Length 1126;
  Best Local Similarity   98.4%;  Pred. No. 2.6e-139;
  Matches  783;  Conservative    4;  Mismatches    3;  Indels    6;  Gaps    6;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 213

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 273

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 333

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 393

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| |||||||||||||||||||||||||||||||||||||| |||||||||
Db        394 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 452

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 512

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 572

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||
Db        573 CATGGGGACATGCACAGCCTGGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGCC 631

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        632 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 691

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        692 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 751

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 811

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        812 GCGTGCCCAACCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAA- 870

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              |:|||||||||||||| |||||||||||||| |:|||||||||||||| ||||||||||
Db        871 GSAGCCGATGTCTGGAG-CTGGGCGTGGCGCTC-TYACCATGCTGGCCGG-CACTACCCC 927

Qy        781 TTCCAGGACTCGGAGC 796
```

```
                    ||| :||||||||||||
Db        928 TTCARGGACTCGGAGC 943


RESULT 15
BX363066
LOCUS       BX363066                944 bp    mRNA    linear   EST 05-MAY-2003
DEFINITION  BX363066 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK008YC10 5-PRIME, mRNA sequence.
ACCESSION   BX363066
VERSION     BX363066.1  GI:30380698
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 944)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DK008BB05QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK008BB05QP1.
FEATURES             Location/Qualifiers
     source          1. .944
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK008YC10"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      157 a    315 c    296 g    168 t      8 others
ORIGIN

  Query Match             66.2%;  Score 711;  DB 13;  Length 944;
  Best Local Similarity   99.2%;  Pred. No. 5.9e-138;
  Matches  744;  Conservative    1;  Mismatches    2;  Indels    3;  Gaps    3;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 256

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 316

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 376

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        377 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGNG 436

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        437 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 495

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 555

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 615

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||
Db        616 CATGGGGACATGCACAGCCTGGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGCC 674

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 734

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTA 794

Qy        601 GAGAACCT-GGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCA 659
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GAGAACCTGGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCA 854

Qy        660 CGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAA 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 CGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAA 914

Qy        720 GGCAGCCGATGTCTGGAGCCTGGGCGTGGC 749
              |||||||||||||||||||||||||||||
Db        915 GGCAGCCGATGTCTGGAGCCTGGGCGTGGC 944


Search completed: January 16, 2004, 16:46:21
Job time : 2414.84 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:09:04 ; Search time 304.409 Seconds
(without alignments)
9524.031 Million cell updates/sec

Title: US-09-909-474D-1_COPY_49_1122
Perfect score: 1074
Sequence: 1 atgcgagccacccctctggc.....gagaagtggttctgtatggc 1074

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1070.8	99.7	1077	22	AAS06709	Polynucleotide seq
2	1069.2	99.6	2059	22	AAI59850	Human polynucleoti
3	1069.2	99.6	2092	22	AAI58064	Human polynucleoti
4	1069.2	99.6	2116	22	AAF30480	Human protein phos
5	1051.6	97.9	1083	24	ABN86479	Human tribbles hom
6	956.4	89.1	1085	21	AAZ61155	cDNA JJ503-KS enco
7	879.2	81.9	2048	21	AAC77866	Human cancer assoc
8	799.2	74.4	972	22	AAH76218	Human kinase PKIN-
9	597.2	55.6	1076	24	ABL39747	Human NS cDNA sequ
10	597.2	55.6	1076	24	ABL39762	Human NS cDNA sequ
c 11	426.2	39.7	2562	16	AAQ89817	DNA encoding the A
12	289.4	26.9	534	23	AAS68600	DNA encoding novel
13	282.2	26.3	396	22	AAS37461	Novel human diagno
14	274.2	25.5	3319	24	ABN86478	Human tribbles hom
15	263.2	24.5	1909	22	AAS45035	cDNA encoding nove
16	251.4	23.4	3280	25	ABX63269	Human cDNA #269 di
17	247	23.0	3324	24	AAS94913	Human DNA sequence
18	236.2	22.0	690	24	ABT09024	Phase-1 Rat CT gen
19	234	21.8	675	24	ABK83550	Human cDNA differe
20	233.4	21.7	942	25	ABX74428	Human cDNA sequenc
21	232.4	21.6	818	25	ACC45125	Human C8FW encodin
22	222.2	20.7	621	25	ACC45126	Human C8FW encodin
23	220.8	20.6	562	21	AAC78131	Human cancer assoc
24	207.6	19.3	1943	22	AAS45223	cDNA encoding nove
c 25	165	15.4	498	24	ABK63084	Rat sequence diffe
26	163.6	15.2	7282	22	AAK86318	Human immune/haema
c 27	158	14.7	562	22	ABA61922	Human foetal liver
c 28	158	14.7	562	22	AAK10231	Human brain expres
c 29	158	14.7	562	22	AAK36130	Human bone marrow
c 30	158	14.7	562	22	AAI41844	Probe #10530 used
c 31	158	14.7	562	23	ABS35823	Human liver single
c 32	158	14.7	562	24	ABS10267	Human genome-deriv
33	158	14.7	25772	22	AAK81332	Human immune/haema
c 34	156	14.5	417	22	ABA74417	Human foetal liver
c 35	156	14.5	417	22	AAK22879	Human brain expres
c 36	156	14.5	417	22	AAK49052	Human bone marrow
c 37	156	14.5	417	22	AAI54879	Probe #23565 used
c 38	156	14.5	417	23	ABS48709	Human liver single
c 39	156	14.5	417	24	ABS22668	Human genome-deriv
40	142.8	13.3	2038	23	ABL29125	Drosophila melanog
c 41	108.2	10.1	6988	23	ABL29124	Drosophila melanog
42	85	7.9	1302	17	AAT38285	Protein kinase cDN
43	85	7.9	1302	20	AAX79647	Human LKB1 coding
44	85	7.9	1302	20	AAX39660	Renal cancer assoc
45	85	7.9	1302	24	ABA03938	Human STK11 coding

ALIGNMENTS

RESULT 1
AAS06709
ID AAS06709 standard; cDNA; 1077 BP.
XX
AC AAS06709;
XX
DT 12-SEP-2001 (first entry)
XX
DE Polynucleotide sequence encoding human protein kinase #9.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Planagan P, Clary D;
XX
DR WPI; 2001-343950/36.

```
DR   P-PSDB; AAU03509.
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections -
XX
PS   Example 1; Figure 1; 433pp; English.
XX
CC   AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC   novel protein kinases have been identified as members of the tyrosine
CC   or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be
CC   used for gene therapy and as DNA probes in diagnostic assays.
CC   The protein kinase polypeptides may be used as antigens in the production
CC   of antibodies against the protein kinases and in assays to identify
CC   modulators of protein kinase expression and activity.
XX
SQ   Sequence 1077 BP; 174 A; 373 C; 325 G; 205 T; 0 other;

  Query Match             99.7%;  Score 1070.8;  DB 22;  Length 1077;
  Best Local Similarity   99.8%;  Pred. No. 6.1e-231;
  Matches 1072;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 360

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074


RESULT 2
AAI59850
ID   AAI59850 standard; cDNA; 2059 BP.
XX
AC   AAI59850;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 3839.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
```

```
DR   P-PSDB; AAM40694.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 3839; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 2059 BP; 400 A; 645 C; 583 G; 431 T; 0 other;

  Query Match            99.6%;  Score 1069.2;  DB 22;  Length 2059;
  Best Local Similarity  99.7%;  Pred. No. 1.6e-230;
  Matches 1071;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 165

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 225

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 285

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 345

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| |||||||||||||||||||||||||||||||||||| ||||||||||
Db        346 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAGGGTGTACCCC 405

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        406 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 465

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 525

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 585

          481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          586 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 645

          541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 705

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 765

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 825

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 885

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 945

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1005

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1065

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1125

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1179


RESULT 3
AAI58064
ID   AAI58064 standard; cDNA; 2092 BP.
XX
AC   AAI58064;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 267.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSB-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM38908.
XX
```

```
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 267; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 2092 BP; 405 A; 658 C; 593 G; 436 T; 0 other;

  Query Match             99.6%;  Score 1069.2;  DB 22;  Length 2092;
  Best Local Similarity   99.7%;  Pred. No. 1.6e-230;
  Matches 1071;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 198

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 258

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 318

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 378

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| |||||||||||||||||||||||||||||||||||| ||||||||||
Db        379 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAGGGTGTACCCC 438

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        439 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 498

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 558

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 618

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 678

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 738

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 798

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 858

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 918

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 978

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1038

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1039 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1098

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1158

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1212

RESULT 4
AAF30480
ID   AAF30480 standard; cDNA; 2116 BP.
XX
AC   AAF30480;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Human protein phosphatase and kinase protein-5 cDNA 1271505CB1.
XX
KW   Protein phosphatase and kinase protein; PPHKP-5; human;
KW   gastrointestinal disorder; immune system disorder;
KW   neurological disorder; cell proliferative disorder; cancer;
KW   diagnosis; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             161..1237
FT                   /*tag=  a
FT   misc_feature    326..385
FT                   /*tag=  b
FT                   /note= "unique fragment"
XX
PN   WO200120004-A2.
XX
PD   22-MAR-2001.
XX
PF   14-SEP-2000; 2000WO-US25515.
XX
PR   15-SEP-1999;   99US-0154141.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Tang YT,  Bandman O,  Hillman JL,  Baughn MR,  Azimzai Y;
PI   Lu DAM;
XX
DR   WPI; 2001-244811/25.
DR   P-PSDB; AAB20326.
XX
PT   Novel human protein phosphatase and kinase proteins for diagnosis,
PT   treatment and prevention of gastrointestinal, immune system,
PT   neurological and cell proliferative disorders -
XX
PS   Claim 5; Page 98; 103pp; English.
```

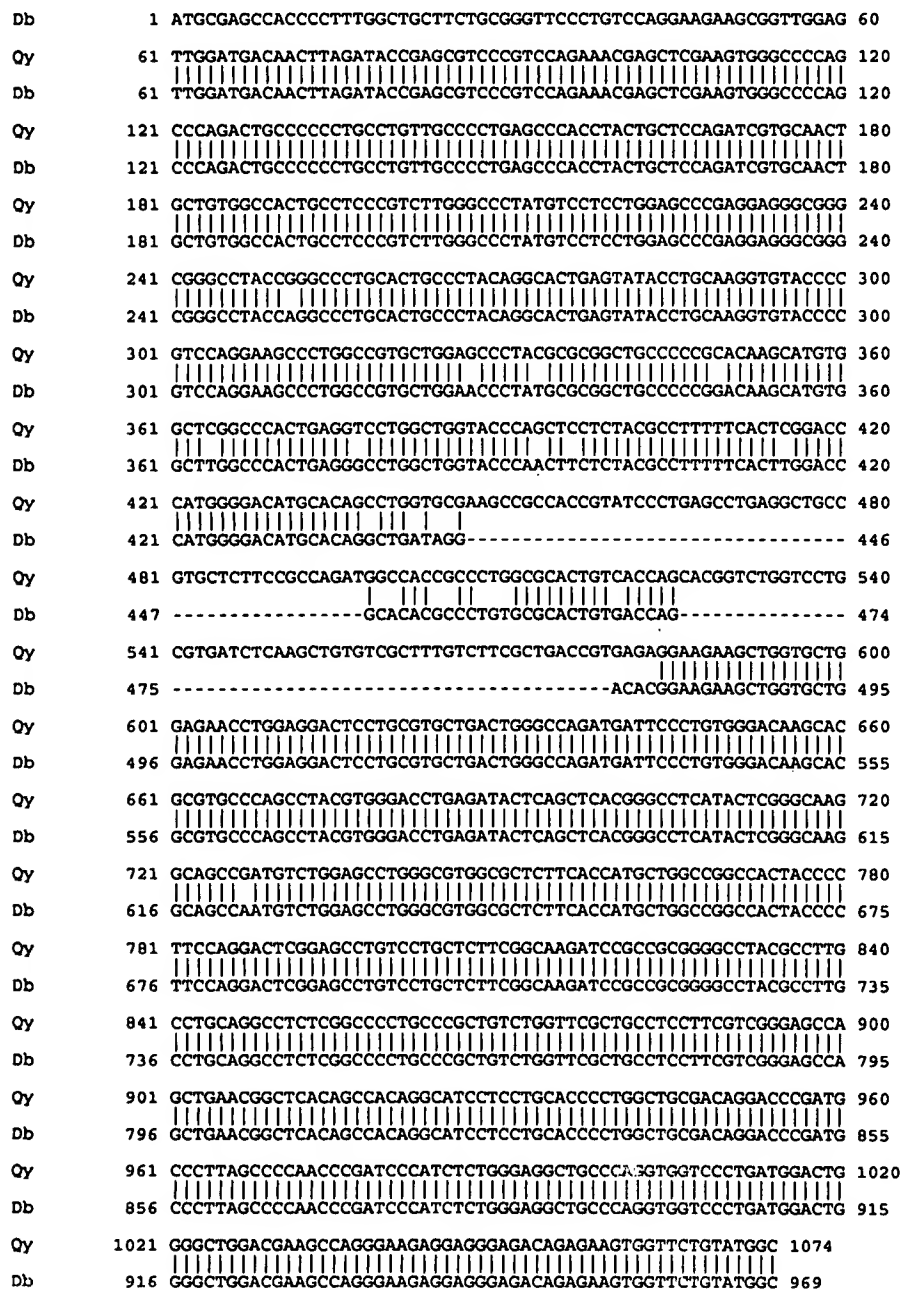
```
XX
CC   The present sequence is that of cDNA encoding novel human
CC   protein phosphatase and kinase protein PPHKP-5 (see AAB20326).
CC   The cDNA was initially identified in Incyte Clone ID No. 1271505CB1,
CC   from a foetal lung tissue library.  The PPHKP-5 gene was localised
CC   to human chromosome 20.  Tissues that express PPHKP-5 (as a
CC   fraction of total tissues expressing PPHKP-5) include reproductive
CC   (0.288), gastrointestinal (0.212) and haematopoietic or immune
CC   (0.192).  Diseases or conditions associated with tissues expressing
CC   PPHKP-5 (as a fraction of total tissues expressing PPHKP-5) include
CC   cancer (0.577), inflammation or trauma (0.327) and cell
CC   proliferation (0.308).  The encoded protein shows homology to rat
CC   kinase.  The invention provides human PPHKP-1 to -11 polypeptides
CC   (see AAB20322-32) and polynucleotides (see AAF30476-86).  It also
CC   provides expression vectors, host cells, antibodies, agonists
CC   and antagonists, as well as methods for diagnosing, treating or
CC   preventing disorders associated with expression of PPHKP, including
CC   gastrointestinal disorders, immune system disorders, neurological
CC   disorders and cell proliferative disorders, including cancer.
XX
SQ   Sequence 2116 BP; 407 A; 663 C; 609 G; 437 T; 0 other;

  Query Match             99.6%;  Score 1069.2;  DB 22;  Length 2116;
  Best Local Similarity   99.7%;  Pred. No. 1.6e-230;
  Matches 1071;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATGCGAGCCACCCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 220

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 280

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 340

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 400

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 460

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 520

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 580

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        581 CATGGGGACATGCACAGCCTGGTGCGAACGCGCCACCGTATCCCTGAGCCTGAGGCTGCC 640

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 700

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 760

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 820

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 880

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 940

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 1000

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1001 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1060

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1061 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1120

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1121 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1180

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1181 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1234


RESULT 5
ABN86479
ID   ABN86479 standard; DNA; 1083 BP.
XX
AC   ABN86479;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human tribbles homologue-3 (htrb-3) polypeptide encoding DNA.
XX
KW   Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic; osteopathic;
KW   opthalamological; cardiant; cytostatic; haemostatic; immunosuppressive;
KW   antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
KW   tumour necrosis factor; TNF; gene; htrb-3; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1083
FT                   /*tag= a
FT                   /product= "htrb-3"
FT                   /note= "tribbles polypeptide"
XX
PN   WO200253743-A2.
XX
PD   11-JUL-2002.
XX
PF   08-JAN-2002; 2002WO-US00070.
XX
PR   08-JAN-2001; 2001US-260294P.
XX
PA   (INTE-) INTERLEUKIN GENETICS INC.
XX
PI   Dower S,  Quanstrom E,  Kiss-Toth E;
XX
DR   WPI; 2002-590635/63.
DR   P-PSDB; ABB80976.
XX
PT   Novel isolated human tribbles homologue-1 polypeptide for inhibiting
PT   AP-1-mediated inflammatory signal in a cell, and activating
PT   ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
PT   -
XX
PS   Example 10; Fig 11A; 131pp; English.
XX
CC   The invention provides an isolated human tribbles homologue-1 (htrb-1,
```

```
CC   also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC   htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC   signal in a cell. The polypeptide employed in the method is preferably
CC   htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or
CC   htrb-3 N C. It is also useful for providing htrb agonist activity for
CC   activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC   signal, an estrogen receptor-mediated gene activation signal, an
CC   fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
CC   in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC   inflammatory signal in a cell such as tumor necrosis factor (TNF)
CC   induced inflammatory signal, or an interleukin induced inflammatory
CC   signal. htrb proteins are useful in screening assays, predictive medicine
CC   and in therapeutics or prophylactics. The htrb proteins are useful for
CC   screening compounds e.g. for treating and/or preventing diseases caused
CC   by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
CC   psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
CC   and cancers. The htrb therapeutics are useful for antagonizing
CC   interleukin-1 dependent disorders of human placenta, intraventricular
CC   hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
CC   and inflammation or autoimmune disorders. The present sequence represents
CC   a DNA encoding the htrb-3 polypeptide.
XX
SQ   Sequence 1083 BP; 176 A; 374 C; 327 G; 206 T; 0 other;

  Query Match             97.9%;  Score 1051.6;  DB 24;  Length 1083;
  Best Local Similarity   99.1%;  Pred. No. 1.3e-226;
  Matches 1070;  Conservative    0;  Mismatches    4;  Indels    6;  Gaps    1;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGGTGCCCCCGCACAAGCATGTG 360

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGC------TGACCGTGAGAGGAAGAAGCTG 594
              |||||||||||||||||||||||||||||||      |||||||||| ||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGACCGTGAGAAGAAGAAGCTG 600

Qy        595 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 660

Qy        655 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 720

Qy        715 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 780

Qy        775 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTAC 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTAC 840

Qy        835 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 900

Qy        895 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 960

Qy        955 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1020

Qy       1015 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1080
```

```
RESULT 6
AAZ61155
ID   AAZ61155 standard; cDNA; 1085 BP.
XX
AC   AAZ61155;
XX
DT   30-MAY-2000  (first entry)
XX
DE   cDNA JJ503-KS encoding domains VIA to XI of a protein kinase.
XX
KW   Kinase activity; molecular weight marker; isoelectric focusing marker;
KW   peptide fragmentation control; cellular signal transduction; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..1081
FT                   /*tag= a
FT                   /note= "partial sequence"
XX
PN   WO200008180-A2.
XX
PD   17-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US17630.
XX
PR   04-AUG-1998;   98US-0095270.
PR   11-SEP-1998;   98US-0099972.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Virca GD,  Bird TA,  Anderson DM,  Marken JS;
XX
DR   WPI; 2000-195584/17.
DR   P-PSDB; AAY69157.
XX
PT   New human kinase polypeptides and polynucleotides used as molecular
PT   weight markers and as controls for peptide fragmentation  -
XX
PS   Claim 1; Page 7; 60pp; English.
XX
CC   The present sequence encodes a partial polypeptide which has kinase
CC   activity. The kinase polynucleotides can be used to express the
CC   polypeptides, and as probes to identify nucleic acids encoding
CC   proteins having kinase activity. The kinase polypeptides and
```

```
CC   fragmented polypeptides are used as molecular weight and isoelectric
CC   focusing markers, and as controls for peptide fragmentation. They also
CC   have a number of therapeutic uses as kinases play a central role in
CC   cellular signal transduction. The polypeptides could also be used to
CC   identify binding partner proteins. The polypeptides can also be used as
CC   a reagent to identify any proteins that the polypeptide regulates, and
CC   proteins with which it might interact. The polypeptides may also be
CC   used for preparation of antibodies. The antibodies can be used in
CC   assays to detect the presence of the protein, and to purify the protein
CC   by immunoaffinity chromatography.
XX
SQ   Sequence 1085 BP; 164 A; 389 C; 338 G; 194 T; 0 other;

  Query Match             89.1%;  Score 956.4;  DB 21;  Length 1085;
  Best Local Similarity   99.9%;  Pred. No. 2.9e-205;
  Matches  957;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 187

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 247

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 307

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 367

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 427

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 487

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 547

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 607

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 667

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 727

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 787

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 847

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        848 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 907

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 967

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        968 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1027

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1085


RESULT 7
AAC77866
ID   AAC77866 standard; cDNA; 2048 BP.
XX
AC   AAC77866;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:260.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43657.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 825-826; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
```

```
XX
SQ   Sequence 2048 BP; 404 A; 619 C; 585 G; 434 T; 6 other;

  Query Match             81.9%;  Score 879.2;  DB 21;  Length 2048;
  Best Local Similarity   92.6%;  Pred. No. 7e-188;
  Matches  994;  Conservative    3;  Mismatches    1;  Indels   76;  Gaps    4;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 215

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 275

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||
Db        276 CCCAGACTGCCCCCCTGCCTGTTGCCCCT------------------------------ 304

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
                                                  ||||||||||||||||||
Db        305 ------------------------------------------GAGCCCGAGGAGGGCGGG 322

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| || |||||||||||||||||||||||||||||| ||||||||||
Db        323 CGGGCCTACCAGG-CCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 380

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        381 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTAYGCGCGGCTGCCCCCGCACAAGCATGTG 440

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 500

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        501 CATGGGGACATGCACAGCCTGGTGCGAAG-CGCCACCGTATCCCTGAGCCTGAGGCTGCC 559

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 619

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 679

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 739

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        740 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCAMGGGCCTCATACTCGGGCAAG 799

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 859

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 919

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 979

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1039

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 CCCTTAGCYCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1099

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1153


RESULT 8
AAH76218
ID   AAH76218 standard; cDNA; 972 BP.
XX
AC   AAH76218;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Human kinase PKIN-10 encoding cDNA.
XX
KW   PKIN; kinase; cytostatic; immunosuppressive; immunostimulant; human;
KW   antiarteriosclerotic; cardiant; gene therapy; antisense therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..972
FT                   /*tag=  a
FT                   /product= "PKIN-10"
XX
PN   WO200160991-A2.
XX
PD   23-AUG-2001.
XX
PF   16-FEB-2001; 2001WO-US05240.
XX
PR   17-FEB-2000; 2000US-0183682.
PR   02-MAR-2000; 2000US-0186559.
PR   09-MAR-2000; 2000US-0188606.
PR   17-MAR-2000; 2000US-0189998.
PR   30-MAR-2000; 2000US-0193851.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Buford N,  Gandhi AR,  Patterson C,  Khan FA,  Yue H;
PI   Hafalia A,  Shih LL,  Tribouley CM,  Yao MG,  Burrill JD,  Marcus GA;
PI   Zingler KA,  Lu DAM,  Bandman O,  Policky JL,  Griffin JA,  Thornton M;
PI   Nguyen DB,  Lal P,  Walsh RT;
XX
DR   WPI; 2001-514771/56.
DR   P-PSDB; AAB85791.
XX
PT   isolated human kinase polypeptides useful in the diagnosis, treatment
PT   and prevention of cancer, immune disorders and disorders affecting
PT   growth and development -
XX
PS   Claim 5; Page 125; 126pp; English.
XX
CC   The invention provides human kinases (PKIN) and polynucleotides encoding
CC   PKIN. The PKIN polypeptides can be expressed using standard recombinant
CC   methodology. The PKIN polypeptides, polynucleotides, modulators and
CC   specific antibodies are useful in the diagnosis, treatment and prevention
CC   of cancer, immune disorders, disorders affecting growth and development,
CC   atherosclerosis, and other cardiovascular diseases, and lipid disorders
CC   and in the assessment of the effects of exogenous compounds on the
CC   expression of nucleic acid sequences of human kinases. The present
CC   sequence represents a cDNA encoding a human PKIN-10 polypeptide.
XX
SQ   Sequence 972 BP; 167 A; 329 C; 294 G; 182 T; 0 other;

  Query Match             74.4%;  Score 799.2;  DB 22;  Length 972;
  Best Local Similarity   87.6%;  Pred. No. 5.5e-170;
  Matches  941;  Conservative    0;  Mismatches   28;  Indels  105;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||| |||||||| |||||||||||||||||||||||||||||||||||
```



RESULT 9
AAL39747

ID ABL39747 standard; cDNA; 1076 BP.
XX
AC ABL39747;
XX
DT 10-MAY-2002 (first entry)
XX
DE Human NS cDNA sequence SEQ ID NO:57.
XX
KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200206315-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-IL00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Freilich S, Bernstein J;
XX
DR WPI; 2002-155037/20.
DR P-PSDB; ABB06093.
XX
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 1; Page 107; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SQ Sequence 1076 BP; 172 A; 352 C; 346 G; 194 T; 12 other;

Query Match 55.6%; Score 597.2; DB 24; Length 1076;
Best Local Similarity 98.1%; Pred. No. 1.2e-124;
Matches 622; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

```
Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 388

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 448

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 508

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 568

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 628

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        629 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 688

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 748

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db        749 CATGGGGACATGCACAGCCTGGTGCGAAGCSSCCACCGTATCCCTGAGCCTGAGGCTGCC 808

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 868

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        869 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG-AGAAGCTGGTGCTG 927

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 634
              ||||  |  |  |||||:||||||||||||||||
Db        928 GAGA--CCTGGAGACTCSTGCGTGCTGACTGGGC 959
```

```
RESULT 10
ABL39762
ID   ABL39762 standard; cDNA; 1076 BP.
XX
AC   ABL39762;
XX
DT   10-MAY-2002  (first entry)
XX
DE   Human NS cDNA sequence SEQ ID NO:72.
XX
KW   Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW   antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW   vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW   anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW   antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW   anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW   gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW   contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW   endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW   rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW   inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW   infertility; cardiovascular disease; coagulation disease; hypertension;
KW   ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW   diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW   gastric ulcer; Alzheimer's disease; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200206315-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-IL00653.
XX
PR   18-JUL-2000; 2000IL-0137345.
PR   15-DEC-2000; 2000IL-0140354.
XX
PA   (COMP-) COMPUGEN LTD.
XX
PI   Mintz L,  Freilich S,  Bernstein J;
XX
DR   WPI; 2002-155037/20.
DR   P-PSDB; ABB06108.
XX
PT   One hundred and twenty eight novel nucleic acid sequences, useful for
PT   treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS   Claim 1; Page 117; 290pp; English.
XX
CC   ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC   encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC   (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC   antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC   vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC   anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC   anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC   immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC   antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC   nootropic and contraceptive activities. The NS can be used in vaccines,
CC   gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC   antibodies from the present invention can be used for treating and
CC   diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC   diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC   cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC   glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC   disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC   disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC   depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC   Alzheimer's disease and as a contraceptive.
XX
SQ   Sequence 1076 BP; 172 A; 352 C; 346 G; 194 T; 12 other;
```

```
  Query Match             55.6%;  Score 597.2;  DB 24;  Length 1076;
  Best Local Similarity   98.1%;  Pred. No. 1.2e-124;
  Matches  622;  Conservative    3;  Mismatches    6;  Indels    3;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 388

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 448

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 508

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 568

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 628

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        629 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 688
```

```
Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 748

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||::||||||||||||||||||||||||||||
Db        749 CATGGGGACATGCACAGCCTGGTGCGAAGCSSCCACCGTATCCCTGAGCCTGAGGCTGCC 808

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 868

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        869 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG-AGAAGCTGGTGCTG 927

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 634
              ||||  |  |  |||||:||||||||||||||||
Db        928 GAGA--CCTGGAGACTCSTGCGTGCTGACTGGGC 959
```

RESULT 11
AAQ89817/c
ID AAQ89817 standard; DNA; 2562 BP.
XX
AC AAQ89817;
XX
DT 25-MAR-2003 (updated)
DT 27-NOV-1995 (first entry)
XX
DE DNA encoding the AUF1 polypeptide.
XX
KW AU-rich element RNA-binding factor-1; proto-oncogene; c-myc; c-myb;
KW c-fos; cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 246..1109
FT /*tag= a
XX
PN WO9511695-A1.
XX
PD 04-MAY-1995.
XX
PF 25-OCT-1994; 94WO-US12273.
XX
PR 29-OCT-1993; 93US-0146421.
XX
PA (UYWA-) UNIV WAKE FOREST.
XX
PI Brewer G;
XX
DR WPI; 1995-178647/23.
DR P-PSDB; AAR74201.
XX
PT AUF1 polypeptide and DNA encoding it - limits the expression of a
PT proto-oncogene that, when over expressed, leads to cancer.
XX
PS Claim 3; Fig 8; 87pp; English.
XX
CC The sequence is that of the DNA encoding AUF1 (AU-rich element RNA-
CC binding factor-1). AUF1 can be used to treat humans with low levels
CC of the AUF1 gene, so limiting the expression of a proto-oncogene (
CC pref. c-myc, c-myb or c-fos) which when over expressed, leads to
CC cancer.
CC See also AAQ89818-20.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2562 BP; 663 A; 608 C; 788 G; 503 T; 0 other;

Query Match 39.7%; Score 426.2; DB 16; Length 2562;
Best Local Similarity 93.8%; Pred. No. 3.4e-86;

Matches 487; Conservative 0; Mismatches 28; Indels 4; Gaps 4;

```
Qy        556 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 615
              || | |||    ||| |     ||   ||||||||||||||||||||||||||||||||
Db       1637 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 1578

Qy        616 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 675
              |||||||||||||||||||||||||||||||| |||||||||||||  |||||||||
Db       1577 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT-GGACAAGCACGCGTCGCCAGCCTAC 1519

Qy        676 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1518 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 1459

Qy        736 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1458 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 1399

Qy        796 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 855
              |||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db       1398 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGCGCTCTCG 1339

Qy        856 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 915
              || |  | ||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1338 CCCTCGCCGCGCTGTCTGGTTCCGTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 1279

Qy        916 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 975
              | |||||  |||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1278 G-CACAGCGATCCTCCTGCACCCCTGGCTGCGACA-GACCCGATGCCCTTAGCCCCAACC 1221

Qy        976 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1035
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1220 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA-CC 1162

Qy       1036 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||
Db       1161 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1123
```

RESULT 12
AAS68600
ID AAS68600 standard; cDNA; 534 BP.
XX
AC AAS68600;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4404.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG04413.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
XX
PS Claim 1; SEQ ID No 4404; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 534 BP; 76 A; 223 C; 150 G; 85 T; 0 other;

Query Match 26.9%; Score 289.4; DB 23; Length 534;
Best Local Similarity 99.7%; Pred. No. 1.3e-55;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 303

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 363

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 423

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 483

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAG 291
              |||||||||| ||||||||||||||||||||||||||||||||||||||
Db        484 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAG 534
```

RESULT 13
AAS37461
ID AAS37461 standard; cDNA; 396 BP.
XX
AC AAS37461;
XX
DT 17-DEC-2001 (first entry)
XX
DE Novel human diagnostic and therapeutic gene #519.
XX
KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
XX
OS Homo sapiens.
XX
PN WO200166753-A2.
XX
PD 13-SEP-2001.
XX
PF 09-MAR-2001; 2001WO-US07787.
XX
PR 09-MAR-2000; 2000US-0188609.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX
DR WPI; 2001-530177/58.
XX
PT New polynucleotides and polypeptides, useful for diagnosis and
PT treatment of breast, lung and colon cancer -
XX
PS Claim 1; Page 727; 1193pp; English.
XX
CC The invention relates to new polynucleotides and polypeptides, useful for
CC diagnosis and treatment of breast, lung and colon cancer. The sequences
CC can be used in detecting differentially expressed genes correlated with a
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample derived from a
CC cell suspected of being cancerous. They can also be used to inhibit
CC tumour growth by modulating expression of a gene product. AAS36943-
CC AAS39338 represent novel human diagnostic and therapeutic coding
CC sequences of the invention.
XX
SQ Sequence 396 BP; 64 A; 138 C; 124 G; 70 T; 0 other;

Query Match 26.3%; Score 282.2; DB 22; Length 396;
Best Local Similarity 97.3%; Pred. No. 5e-54;
Matches 287; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        101 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGATGGAG 160

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        161 TTGGATGACAACTTAGATACCGAGCGTTCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 220

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 280

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 340

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGT 295
              ||| |||||| |||||||||||||||||||||    ||||||||||||||||
Db        341 CGGTCCTACCAGGCCCTGCACTGCCCTACAGGCACTTGAGATACCTGCAAGGTGT 395
```

RESULT 14
ABN86478
ID ABN86478 standard; cDNA; 3319 BP.
XX
AC ABN86478;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human tribbles homologue-1 (htrb-1) polypeptide encoding cDNA.
XX
KW Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic; osteopathic;
KW opthalamological; cardiant; cytostatic; haemostatic; immunosuppressive;
KW antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
KW tumour necrosis factor; TNF; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

```
PT   CDS             282..1400
PT                   /*tag=  a
PT                   /product= "htrb-1"
PT                   /note= "tribbles polypeptide"
XX
PN   WO200253743-A2.
XX
PD   11-JUL-2002.
XX
PF   08-JAN-2002; 2002WO-US00070.
XX
PR   08-JAN-2001; 2001US-260294P.
XX
PA   (INTE-) INTERLEUKIN GENETICS INC.
XX
PI   Dower S,  Quanstrom E,  Kiss-Toth E;
XX
DR   WPI; 2002-590635/63.
DR   P-PSDB; ABB80975.
XX
PT   Novel isolated human tribbles homologue-1 polypeptide for inhibiting
PT   AP-1-mediated inflammatory signal in a cell, and activating
PT   ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
PT    -
XX
PS   Claim 5; Fig 10A; 131pp; English.
XX
CC   The invention provides an isolated human tribbles homologue-1 (htrb-1,
CC   also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC   htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC   signal in a cell. The polypeptide employed in the method is preferably
CC   htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or
CC   htrb-3 N C. It is also useful for providing htrb agonist activity for
CC   activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC   signal, an estrogen receptor-mediated gene activation signal, an
CC   fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
CC   in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC   inflammatory signal in a cell such as tumor necrosis factor (TNF)
CC   induced inflammatory signal, or an interleukin induced inflammatory
CC   signal. htrb proteins are useful in screening assays, predictive medicine
CC   and in therapeutics or prophylactics. The htrb proteins are useful for
CC   screening compounds e.g. for treating and/or preventing diseases caused
CC   by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
CC   psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
CC   and cancers. The htrb therapeutics are useful for antagonizing
CC   interleukin-1 dependent disorders of human placenta, intraventricular
CC   hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
CC   and inflammation or autoimmune disorders. The present sequence represents
CC   a cDNA encoding the htrb-1 polypeptide.
XX
SQ   Sequence 3319 BP; 782 A; 909 C; 800 G; 827 T; 1 other;

  Query Match             25.5%;  Score 274.2;  DB 24;  Length 3319;
  Best Local Similarity   60.3%;  Pred. No. 4.6e-52;
  Matches  453;  Conservative    0;  Mismatches  298;  Indels    0;  Gaps    0;

Qy        194 CCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGG 253
Db        544 CCAGCCGCATCGCCGACTACCTGCTGCTGCCCCTAGCCGAGCGCGAGCATGTGTCCCGGG 603

Qy        254 CCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCC 313
Db        604 CGCTGTGCATCCACACTGGACGCGAGCTGCGCTGCAAGGTGTTTCCCATTAAACACTACC 663

Qy        314 TGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTG 373
Db        664 AGGACAAAATCAGGCCTTACATCCAGCTGCCATCGCACAGCAACATTACTGGCATTGTGG 723

Qy        374 AGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGC 433
Db        724 AAGTGATCCTTGGGGAAACCAAGGCCTATGTCTTCTTTGAGAAGAGCTTTGGGGACATGC 783

Qy        434 ACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCC 493
Db        784 ACTCCTATGTGCGAAGCCGGAAGAGGCTGCGGGAAGAGGAAGCCGCCCGGCTCTTCAAGC 843

Qy        494 AGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGC 553
Db        844 AGATTGTCTCCGCCGTCGCCCACTGCCACCAGTCAGCCATCGTGCTGGGGGACCTGAAGC 903

Qy        554 TGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGG 613
Db        904 TTAGGAAGTTCGTCTTCTCCACGGAGGAGAGAACCCAGCTTAGACTAGAAAGTCTAGAAG 963

Qy        614 ACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCT 673
Db        964 ACACACACATAATGAAGGGGGAAGATGATGCTTTGTCAGACAAACATGGCTGCCCAGCCT 1023

Qy        674 ACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCT 733
Db       1024 ACGTGAGCCCTGAGATCCTCAACACCACTGGGACCTACTCCGGAAAGGCTGCGGACGTTT 1083

Qy        734 GGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGG 793
Db       1084 GGAGCCTGGGGGTGATGCTCTACACCCTTCTGGTTGGACGATACCCCTTCCATGACTCAG 1143

Qy        794 AGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCT 853
Db       1144 ACCCCAGTGCCCTTTTCTCCAAAATTCGGCGTGGACAGTTCTGCATTCCTGAGCACATTT 1203

Qy        854 CGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCA 913
Db       1204 CCCCCAAAGCCAGGTGCCTCATTCGCAGCCTCTTGAGACGGGAGCCCTCCGAGAGACTCA 1263

Qy        914 CAGCCACAGGCATCCTCCTGCACCCCTGGCT 944
Db       1264 CTGCCCCCGAGATCCTACTGCACCCCTGGTT 1294


RESULT 15
AAS45035
ID   AAS45035 standard; cDNA; 1909 BP.
XX
AC   AAS45035;
XX
DT   18-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human secretory protein, Seq ID No 116.
XX
KW   Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW   ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW   transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW   amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW   ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW   gut protection; lung; liver fibrosis; immune deficiency; infection;
KW   severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW   multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW   fertility; analgesic; pain; antigen; ss.
XX
OS   Homo sapiens.
XX
PN   WO200166689-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US04942.
XX
PR   07-MAR-2000; 2000US-0519705.
PR   19-MAY-2000; 2000US-0574454.
PR   17-JUN-2000; 2000US-0596193.
PR   14-JUL-2000; 2000US-0616847.
PR   19-SEP-2000; 2000US-0665363.
PR   20-OCT-2000; 2000US-0693267.
XX
PA   (HYSE-) HYSEQ INC.
XX
```

PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
DR WPI; 2001-589934/66.
DR P-PSDB; AAU28135.
XX
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS Claim 1; SEQ ID No 116; 107pp; English.
XX
CC The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAS44920-AAS45295 represent novel human secreted protein
CC coding sequences of the invention.
XX
SQ Sequence 1909 BP; 472 A; 535 C; 435 G; 467 T; 0 other;

Query Match 24.5%; Score 263.2; DB 22; Length 1909;
Best Local Similarity 60.7%; Pred. No. 1.2e-49;
Matches 447; Conservative 0; Mismatches 288; Indels 1; Gaps 1;

```
Qy        210 CTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTAC 269
              |||  | || |||   |  |  ||| ||| ||  |  | ||||| |||  |  ||  ||
Db        103 CTACCTGCTGCTGCCCCTAGCCGAGCGCGAGCATGTGTCCCGGGCGCTGTGCATCCACAC 162

Qy        270 AGGCACTGAGTATAC-CTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGC 328
              || ||   |  | | ||||||||||| ||| |  |  |   || || |    |   ||
Db        163 TGGGACGCCGAGTGCGCTGCAAGGTGTTTCCCATTAAACACTACCAGGACAAAATCAGGC 222

Qy        329 CCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTA 388
              | |||   | ||||||  ||||||   |  | || |       || ||  |   |||
Db        223 CTTACATCCAGCTGCCATCGCACAGCAACATTACTGGCATTGTGGAAGTGATCCTTGGGG 282

Qy        389 CCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAA 448
                  ||| | ||| ||       |  |  ||||||||||||  |   |||||||
Db        283 AAACCAAGGCCTATGTCTTCTTTGAGAAGGACTTTGGGGACATGCACTCCTATGTGCGAA 342

Qy        449 GCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCC 508
              ||||  |  |  | |  ||   || || |||  |||||||   |||||  | |  |||||
Db        343 GCCGGAAGAGGCTGCGGGAAGAGGAAGCCGCCCGGCTCTTCAAGCAGATTGTCTCCGCCG 402

Qy        509 TGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCT 568
```

```
              | || |||||  ||||||   |   | || ||| | || || |||||  |    || ||||
Db        403 TCGCCCACTGCCACCAGTCAGCCATCGTGCTGGGGGACCTGAAGCTTAGGAAGTTCGTCT 462

Qy        569 TCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGA 628
              || |       |||||  |  ||||    || || |  || || ||| |   | |  |||
Db        463 TCTCCACGGAGGAGAGAACCCAGCTTAGACTAGAAAGTCTAGAAGACACACACATAATGA 522

Qy        629 CTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGA 688
                |||  |||||||  |  |||  |||||  || |  |||||||||||||||  | |||||||
Db        523 AGGGGGAAGATGATGCTTTGTCAGACAAACATGGCTGCCCAGCCTACGTGAGCCCTGAGA 582

Qy        689 TACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGG 748
              | |||| | |     |   | ||||| || ||||| || || || ||||||||||| |||
Db        583 TCCTCAACACCACTGGGACCTACTCCGGAAAGGCTGCGGACGTTTGGAGCCTGGGGGTGA 642

Qy        749 CGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCT 808
              |||||  |||| | ||||  || |  |||||||||||  |||||  || ||       || |
Db        643 TGCTCTACACCCTTCTGGTTGGACGATACCCCTTCCATGACTCAGACCCCAGTGCCCTTT 702

Qy        809 TCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCT 868
              ||  ||| || || || ||     | |  | | ||||    | | ||  ||  ||| | |
Db        703 TCTCCAAAATTCGGCGTGGACAGTTCTGCATTCCTGAGCACATTTCCCCCAAAGCCAGGT 762

Qy        869 GTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCC 928
              | ||  |||||  |||||  |  |||||||||  |  ||  | |||||  ||| | |  ||||
Db        763 GCCTCATTCGCAGCCTCTTGAGACGGGAGCCCTCCGAGAGACTCACTGCCCCCGAGATCC 822

Qy        929 TCCTGCACCCCTGGCT 944
              | ||||||||||||  |
Db        823 TACTGCACCCCTGGTT 838
```

Search completed: January 16, 2004, 11:33:02
Job time : 311.409 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:11:12 ; Search time 4024.84 Seconds
(without alignments)
10916.435 Million cell updates/sec

Title: US-09-909-474D-1_COPY_49_1122
Perfect score: 1074
Sequence: 1 atgcgagccacccctctggc.....gagaagtggttctgtatggc 1074

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1070.8	99.7	1077	6	AX166518	AX166518 Sequence
2		1070.8	99.7	2121	9	BC019363	BC019363 Homo sapi
3		1070.8	99.7	2283	9	BC027484	BC027484 Homo sapi
4		1070.8	99.7	2488	9	AY247738	AY247738 Homo sapi
5		1069.2	99.6	2116	6	AX099934	AX099934 Sequence
6		1066	99.3	2257	9	AK026945	AK026945 Homo sapi
7		1051.6	97.9	1083	6	AX572896	AX572896 Sequence
8		1051.6	97.9	1083	9	AF250311	AF250311 Homo sapi
9		799.2	74.4	972	6	AX224734	AX224734 Sequence
10		656.6	61.1	1256	10	AF358868	AF358868 Mus muscu
11		611.2	56.9	2004	10	AB020967	AB020967 Rattus sp
12		597.2	55.6	1076	6	AX364906	AX364906 Sequence
13		597.2	55.6	1076	6	AX364921	AX364921 Sequence
14		512.4	47.7	1969	10	BC012955	BC012955 Mus muscu
15		495	46.1	153170	9	HS1103G7	AL034548 Human DNA
16	c	426.2	39.7	2559	6	AR274998	AR274998 Sequence
17	c	426.2	39.7	2559	9	HSU02019	U02019 Human AU-ri
18	c	426.2	39.7	2562	6	I24432	I24432 Sequence 4
19	c	297	27.7	75131	2	AC023022	AC023022 Homo sapi
20	c	296.6	27.6	191503	2	AL928568	AL928568 Mus muscu
21	c	286.8	26.7	235947	2	AC133773	AC133773 Rattus no
22	c	286.8	26.7	282849	2	AC132555	AC132555 Rattus no
23		282.2	26.3	396	6	AX245589	AX245589 Sequence
24		274.2	25.5	3319	6	AX572894	AX572894 Sequence
25		274.2	25.5	3319	9	AF250310	AF250310 Homo sapi
26		272.6	25.4	3317	9	AF205437	AF205437 Homo sapi
27		272.6	25.4	3338	9	HSM803696	AL832388 Homo sapi
28		261.4	24.3	2514	10	BC006800	BC006800 Mus muscu
29		258.2	24.0	1244	10	AF358866	AF358866 Mus muscu
30		254	23.6	1682	4	AY247741	AY247741 Bos tauru
31		253	23.6	2156	9	BC002637	BC002637 Homo sapi
32		253	23.6	3062	9	AY245544	AY245544 Homo sapi
33		253	23.6	4221	9	D87119	D87119 Homo sapien
34		250.6	23.3	1102	10	AF205438	AF205438 Rattus no
35		247	23.0	3324	6	AX281759	AX281759 Sequence
36		244.4	22.8	1032	4	CFC5FW	X99144 C.familiari
37		241.2	22.5	1496	10	AF358867	AF358867 Mus muscu
38		241.2	22.5	3244	10	BC034338	BC034338 Mus muscu
39		236.2	22.0	690	6	AX525590	AX525590 Sequence
40		234	21.8	675	9	HSC8FW	AJ000480 Homo sapi
41		222.2	20.7	2667	10	BC027159	BC027159 Mus muscu
42		217.8	20.3	1747	5	AY247742	AY247742 Gallus ga
43		210.8	19.6	2136	5	AY254200	AY254200 Xenopus l
44	c	165	15.4	498	6	AX401315	AX401315 Sequence
45		163.6	15.2	123529	9	AF179296	AF179296 Homo sapi

ALIGNMENTS

RESULT 1
AX166518
LOCUS AX166518 1077 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 9 from Patent WO0138503.
ACCESSION AX166518
VERSION AX166518.1 GI:14546863
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Planagan,P. and Clary,D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes

```
  JOURNAL   Patent: WO 0138503-A 9 31-MAY-2001;
            Sugen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1077
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      174 a    373 c    325 g    205 t
ORIGIN

  Query Match             99.7%;  Score 1070.8;  DB 6;  Length 1077;
  Best Local Similarity   99.8%;  Pred. No. 6.5e-188;
  Matches 1072;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 360

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074


RESULT 2
BC019363
LOCUS       BC019363                2121 bp    mRNA    linear   PRI 06-JUN-2003
DEFINITION  Homo sapiens chromosome 20 open reading frame 97, mRNA (cDNA clone
            MGC:854 IMAGE:3528490), complete cds.
ACCESSION   BC019363
VERSION     BC019363.1  GI:17939414
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2121)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2121)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
```

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 8 Row: g Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11056039.

```
FEATURES             Location/Qualifiers
     source          1. .2121
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:854 IMAGE:3528490"
                     /tissue_type="Muscle, rhabdomyosarcoma"
                     /clone_lib="NIH_MGC_17"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .2121
                     /gene="C20orf97"
                     /note="synonyms: dJ1103G7.3, SKIP3, TRB3"
                     /db_xref="LocusID:57761"
     CDS             108. .1184
                     /codon_start=1
                     /product="chromosome 20 open reading frame 97"
                     /protein_id="AAH19363.1"
                     /db_xref="GI:17939415"
                     /db_xref="LocusID:57761"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQ
                     EALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                     VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD
                     KHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG
                     AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ
                     VVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      459 a    645 c    584 g    433 t
ORIGIN
```

Query Match 99.7%; Score 1070.8; DB 9; Length 2121;
Best Local Similarity 99.8%; Pred. No. 5.7e-188;
Matches 1072; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 167

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 227

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 287

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 347

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 407

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        408 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 467

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 527

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 587

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 647

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 707

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        708 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 767

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 827

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        828 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 887

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        888 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 947

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        948 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1007

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1067

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1127

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1181
```

```
RESULT 3
BC027484
LOCUS       BC027484                2283 bp    mRNA    linear   PRI 08-APR-2002
DEFINITION  Homo sapiens, chromosome 20 open reading frame 97, clone MGC:34909
            IMAGE:5104452, mRNA, complete cds.
ACCESSION   BC027484
VERSION     BC027484.1  GI:20071610
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2283)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
```

```
           Richards, S., Gibbs, R.A.

           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series: IRAK Plate: 52 Row: f Column: 11.
FEATURES             Location/Qualifiers
     source          1. .2283
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="LocusID:57761"
                     /db_xref="taxon:9606"
                     /clone="MGC:34909 IMAGE:5104452"
                     /tissue_type="Cervix, carcinoma"
                     /clone_lib="NIH_MGC_12"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             295. .1371
                     /codon_start=1
                     /product="chromosome 20 open reading frame 97"
                     /protein_id="AAH27484.1"
                     /db_xref="GI:20071611"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYRALHCPTGTEYTCKVYPVQ
                     EALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                     VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD
                     KHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG
                     AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ
                     VVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      463 a    702 c    656 g    462 t
ORIGIN

  Query Match             99.7%;  Score 1070.8;  DB 9;  Length 2283;
  Best Local Similarity   99.8%;  Pred. No. 5.6e-188;
  Matches 1072;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATGCGAGCCACCCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 354

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 414

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 474

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 534

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 594

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 654

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 714

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 774

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 834

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTA 894

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 954

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 1014

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 1074

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 1134

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1194

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1254

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 CCCTTAGCTCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1314

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1368


RESULT 4
AY247738
LOCUS       AY247738               2488 bp    mRNA    linear   PRI 20-APR-2003
DEFINITION  Homo sapiens TRB3 protein mRNA, complete cds.
ACCESSION   AY247738
VERSION     AY247738.1  GI:30025661
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2488)
  AUTHORS   Shan,Y.X. and Yu,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-2003) School of Life Science, Institute of
            Genetics, Fudan University, Handan Road 220, Shanghai 200433, PR
            China
FEATURES             Location/Qualifiers
     source          1. .2488
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             541. .1617
                     /codon_start=1
                     /product="TRB3 protein"
                     /protein_id="AAP04407.1"
                     /db_xref="GI:30025662"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQ
                     EALAVLEPYARLPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                     VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD
                     KHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG
                     AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ
                     VVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      478 a    768 c    746 g    496 t
ORIGIN

  Query Match             99.7%;  Score 1070.8;  DB 9;  Length 2488;
```

```
  Best Local Similarity   99.8%;  Pred. No. 5.5e-188;
  Matches 1072;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 600

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 660

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 720

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 780

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 840

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        841 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 900

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 960

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 1020

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 1080

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 1140

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 1200

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 1260

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 1320

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 1380

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1440

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1500

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1560

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1614


RESULT 5
AX099934
LOCUS       AX099934               2116 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 16 from Patent WO0120004.
ACCESSION   AX099934
VERSION     AX099934.1  GI:13538944
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yue,H., Tang,Y.T., Bandman,O., Hillman,J.L., Baughn,M.R.,
            Azimzai,Y. and Lu,D.A.
  TITLE     Protein phosphatase and kinase proteins
  JOURNAL   Patent: WO 0120004-A 16 22-MAR-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2116
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID NO: 1271505CB1"
BASE COUNT      407 a    663 c    609 g    437 t
ORIGIN

  Query Match             99.6%;  Score 1069.2;  DB 6;  Length 2116;
  Best Local Similarity   99.7%;  Pred. No. 1.1e-187;
  Matches 1071;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 220

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 280

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 340

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 400

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 460

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 520

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 580

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        581 CATGGGGACATGCACAGCCTGGTGCGAACGCGCCACCGTATCCCTGAGCCTGAGGCTGCC 640

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 700

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 760
```

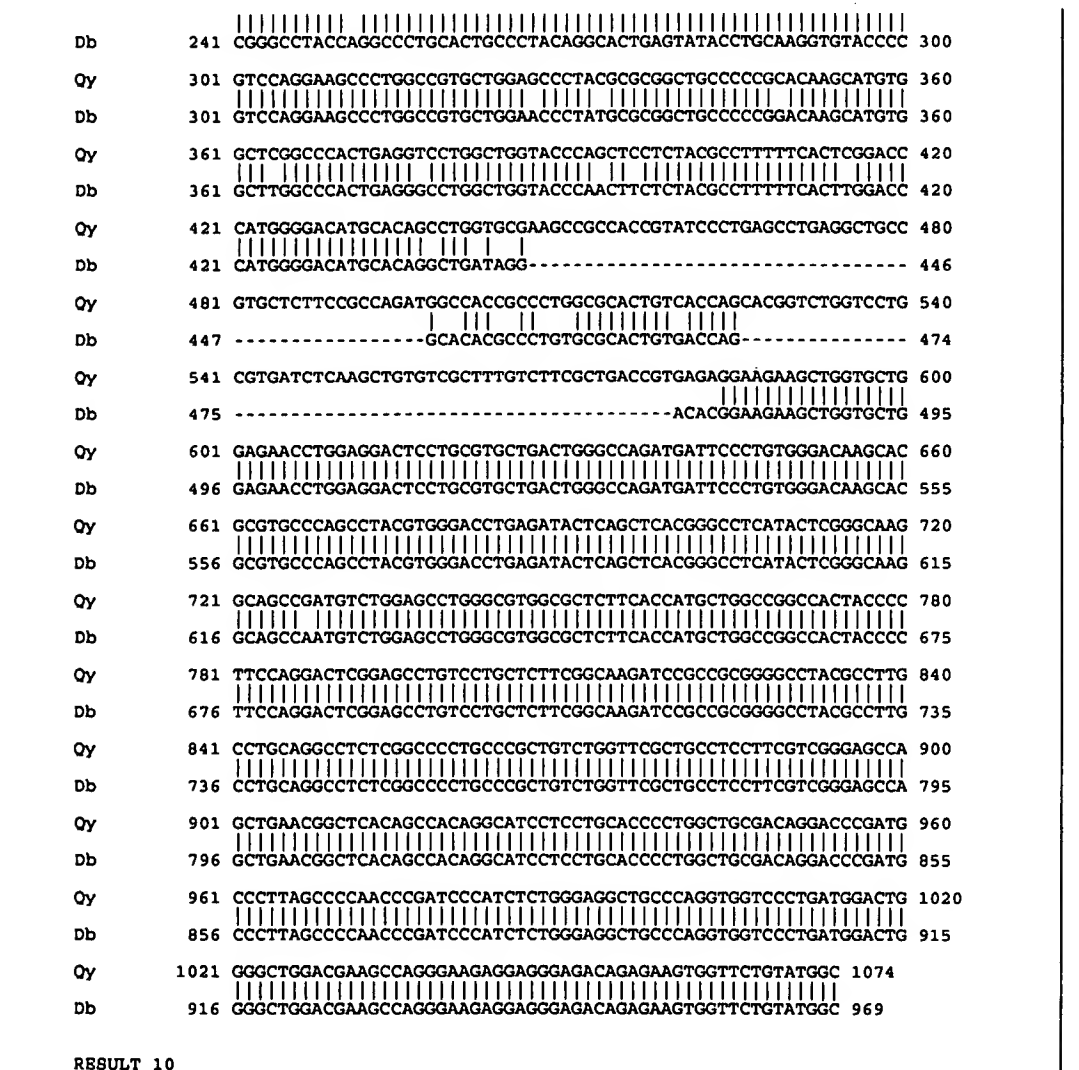
```
Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 820

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 880

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 940

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 1000

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1001 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1060

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1061 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1120

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1121 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1180

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1181 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1234


RESULT 6
AK026945
LOCUS       AK026945               2257 bp    mRNA    linear   PRI 29-SEP-2000
DEFINITION  Homo sapiens cDNA: FLJ23292 fis, clone HEP10334.
ACCESSION   AK026945
VERSION     AK026945.1  GI:10439923
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
            Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
            Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2257)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .2257
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEP10334"
                     /cell_line="HepG2"
                     /cell_type="hepatoma"
                     /clone_lib="HEP"
                     /note="cloning vector pME18SFL3"
     CDS             295. .1371
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB15597.1"
                     /db_xref="GI:10439924"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEGGRAYQALHCPTGTEYTCKVYPVQ
                     EAPAVLEPYARLPPHKHVARPTEVLAGTQLLYAPFTRTHGDMHSLVRSRHRIPEPEAA
                     VLFRQMATALAHCHQHGLVLRDLKLCRFVFADRERKKLVLENLEDSCVLTGPDDSLWD
                     KHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIRRG
                     AYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEAAQ
                     VVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      435 a    705 c    659 g    458 t
ORIGIN

  Query Match             99.3%;  Score 1066;  DB 9;  Length 2257;
  Best Local Similarity   99.5%;  Pred. No. 4.3e-187;
  Matches 1069;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATGCGAGCCACCCCTCTAGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 354

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 414

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 474

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 534

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 594

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||||||
Db        595 GTCCAGGAAGCCCCGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 654

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 714

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 774

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 834

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 894

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 954

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 1014

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 1074
```

```
Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGCGGGGCCTACGCCTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGCGGGGCCTACGCCTTG 1134

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 1194

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1254

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1255 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGTCTG 1314

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1368
```

```
RESULT 7
AX572896
LOCUS       AX572896               1083 bp    DNA     linear   PAT 29-NOV-2002
DEFINITION  Sequence 3 from Patent WO02053743.
ACCESSION   AX572896
VERSION     AX572896.1  GI:26004968
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dower,S. and Quanstrom,E.
  TITLE     Mammalian tribbles signaling pathways and methods and reagents
            related thereto
  JOURNAL   Patent: WO 02053743-A 3 11-JUL-2002;
            Interleukin Genetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1083
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      176 a    374 c    327 g    206 t
ORIGIN

  Query Match             97.9%;  Score 1051.6;  DB 6;  Length 1083;
  Best Local Similarity   99.1%;  Pred. No. 2.3e-184;
  Matches 1070;  Conservative    0;  Mismatches    4;  Indels    6;  Gaps    1;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGGTGCCCCCGCACAAGCATGTG 360

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGC------TGACCGTGAGAGGAAGAAGCTG 594
              |||||||||||||||||||||||||||||||      |||||||||| ||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGACCGTGAGAAGAAGAAGCTG 600

Qy        595 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 660

Qy        655 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 720

Qy        715 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 780

Qy        775 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTAC 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTAC 840

Qy        835 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 900

Qy        895 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 960

Qy        955 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1020

Qy       1015 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1080
```

```
RESULT 8
AF250311
LOCUS       AF250311               1083 bp    mRNA    linear   PRI 01-JUN-2001
DEFINITION  Homo sapiens SKIP3 mRNA, complete cds.
ACCESSION   AF250311
VERSION     AF250311.1  GI:14276268
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1083)
  AUTHORS   Kiss-Toth,E., Wyllie,D.H., Qwarnstrom,E.E. and Dower,S.K.
  TITLE     Identification of pro-inflammatory cytokine signalling network
            components by transcription expression screening
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1083)
  AUTHORS   Kiss-Toth,E., Wyllie,D.H., Qwarnstrom,E.E. and Dower,S.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAR-2000) Division of Molecular & Genetic Medicine,
            University of Sheffield, Royal Hallamshire Hospital, Floor M,
            Sheffield S10-2JF, UK
```

```
FEATURES             Location/Qualifiers
     source          1. .1083
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             1. .1083
                     /codon_start=1
                     /product="SKIP3"
                     /protein_id="AAK58175.1"
                     /db_xref="GI:14276269"
                     /translation="MRATPLAAPAGSLSRKKRLELDDNLDTERPVQKRARSGPQPRLP
                     PCLLPLSPPTAPDRATAVATASRLGPYVLLEPEEGGRAYQALHCPTGTEYTCKVYPVQ
                     EALAVLEPYARVPPHKHVARPTEVLAGTQLLYAFFTRTHGDMHSLVRSRHRIPEPEAA
                     VLPRQMATALAHCHQHGLVLRDLKLCRFVFADRDREKKKLVLENLEDSCVLTGPDDSL
                     WDKHACPAYVGPEILSSRASYSGKAADVWSLGVALFTMLAGHYPFQDSEPVLLFGKIR
                     RGAYALPAGLSAPARCLVRCLLRREPAERLTATGILLHPWLRQDPMPLAPTRSHLWEA
                     AQVVPDGLGLDEAREEEGDREVVLYG"
BASE COUNT      176 a    374 c    327 g    206 t
ORIGIN

  Query Match           97.9%;  Score 1051.6;  DB 9;  Length 1083;
  Best Local Similarity   99.1%;  Pred. No. 2.3e-184;
  Matches 1070;  Conservative    0;  Mismatches    4;  Indels    6;  Gaps    1;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGGTGCCCCCGCACAAGCATGTG 360

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGC------TGACCGTGAGAGGAAGAAGCTG 594
              |||||||||||||||||||||||||||||||      |||||||||| ||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGACCGTGAGAAGAAGAAGCTG 600

Qy        595 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 660

Qy        655 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 720

Qy        715 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 780

Qy        775 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTAC 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTAC 840

Qy        835 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 900

Qy        895 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 960

Qy        955 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1020

Qy       1015 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1080


RESULT 9
AX224734
LOCUS       AX224734                 972 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 21 from Patent WO0160991.
ACCESSION   AX224734
VERSION     AX224734.1  GI:15554836
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Buford,N., Gandhi,A.R., Patterson,C., Khan,F.A., Yue,H.,
            Hafalia,A., Shih,L.L., Tribouley,C.M., Yao,M.G., Burrill,J.D.,
            Marcus,G.A., Zingler,K.A., Lu,D.A., Bandman,O., Policky,J.L.,
            Griffin,J.A., Thornton,M., Nguyen,D.B. and Walsh,R.T.
  TITLE     Human kinases
  JOURNAL   Patent: WO 0160991-A 21 23-AUG-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .972
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472328CB1"
BASE COUNT      167 a    329 c    294 g    182 t
ORIGIN

  Query Match           74.4%;  Score 799.2;  DB 6;  Length 972;
  Best Local Similarity   87.6%;  Pred. No. 9.6e-138;
  Matches  941;  Conservative    0;  Mismatches   28;  Indels  105;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTTTGGCTGCTTCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
```

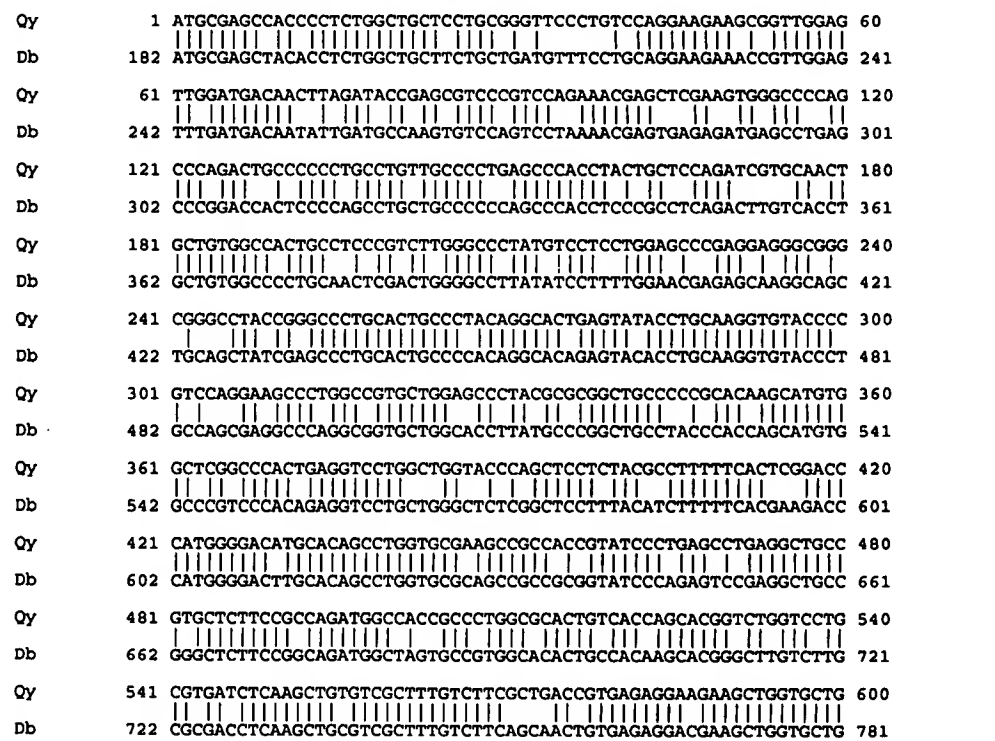


RESULT 10
AF358868
LOCUS AF358868 1256 bp mRNA linear ROD 01-JUN-2002
DEFINITION Mus musculus TRB-3 mRNA, complete cds.
ACCESSION AF358868
VERSION AF358868.1 GI:21304713
KEYWORDS .
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1256)
AUTHORS Kiss-Toth,E., Dempsey,C., Jozsa,V., Caunt,J., Oxley,K.M., Bagstaff,S.M., Wyllie,D.H., Harte,M., O'Neill,L.A.J., Qwarnstrom,E.E. and Dower,S.K.
TITLE Mammalian homologs of Drosophila tribbles (htrb) control mitogen activated protein kinase signaling

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1256)
AUTHORS Kiss-Toth,E., Dempsey,C., Jozsa,V., Caunt,J., Oxley,K.M., Bagstaff,S.M., Wyllie,D.H., Harte,M., O'Neill,L.A.J., Qwarnstrom,E.E. and Dower,S.K.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2001) Division of Genomic Medicine, University of Sheffield, Royal Hallamshire Hospital, Floor M, Glossop Road, Sheffield S10-2JF, UK

FEATURES Location/Qualifiers
source 1. .1256
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
CDS 182. .1246
/note="mTRB-3"
/codon_start=1
/product="TRB-3"
/protein_id="AAM45476.1"
/db_xref="GI:21304714"
/translation="MRATPLAASADVSCRKKPLEFDDNIDAKCPVLKRVRDEPEPGPL
PSLLPPSPPPASDLSPAVAPATRLGPYILLEREQGSCSYRALHCPTGTEYTCKVYPAS
EAQAVLAPYARLPTHQHVARPTEVLLGSRLLYIFFTKTHGDLHSLVRSRRGIPESEAA
GLFRQMASAVAHCHKHGLVLRDLKLRRFVFSNCERTKLVLENLEDACVMTGSDDSLWD
THACPAYVGPEILSSRPSYSGKAADVWSLGVALFTMLAGRYPFHDSEPVLLFGKIRXG
TFALPEGLSAPARCLIRCLLRKEPSERLVALGILLHPWLREDHGRVSPPQSDRREMDQ
VVPDGPQLEEAEEGEVGLYG"
BASE COUNT 231 a 390 c 385 g 249 t 1 others
ORIGIN

Query Match 61.1%; Score 656.6; DB 10; Length 1256;
Best Local Similarity 76.6%; Pred. No. 2e-111;
Matches 803; Conservative 0; Mismatches 245; Indels 0; Gaps 0;



```
Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              |||||||||||  ||  ||||||||  |||||||  |||||||  ||  ||||||||||  |||
Db        782 GAGAACCTGGAAGATGCCTGCGTGATGACTGGATCAGATGACTCTCTGTGGGACACGCAT 841

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||  ||||||||||||||  ||||||||||||||  |||  |  ||  |||||  |||||
Db        842 GCGTGCCCTGCCTACGTGGGACCAGAGATACTCAGCTCCCGGCCATCCTACTCTGGCAAA 901

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||  ||  ||||||||||||||||  ||||||||||||||||||||||||  |||    |||||
Db        902 GCGGCTGATGTCTGGAGCCTGGGTGTGGCGCTCTTCACCATGCTGGCTGGCAGATACCCA 961

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              |||||  |||||  |||||  ||||||||||  ||||||||||    |||  |||   |||  ||
Db        962 TTCCACGACTCTGAGCCAGTCCTGCTCTTTGGCAAGATCCGTANAGGGACCTTTGCCCTG 1021

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||  |||||  ||  |||||  ||||||||||||  |  |||||  |||||  ||   |||  ||
Db       1022 CCTGAGGGCCTATCAGCCCCAGCCCGCTGTCTGATCCGCTGTCTCCTCCGCAAGGAACCT 1081

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              |  ||  ||  ||    |||   ||  ||||||  ||||  ||||||  ||  ||  ||||  |
Db       1082 TCAGAGCGACTTGTGGCCCTGGGAATCCTCTTGCATCCCTGGTTGAGAGAGGATCACGGC 1141

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              |  |  |  ||  |  ||  |  |   |||||   |  ||||||||||||  |||||  |
Db       1142 CGAGTCTCTCCTCCACAGTCTGACCGAAGGGAGATGGACCAGGTGGTCCCAGATGGGCCA 1201

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGG 1048
                |||||||  ||  ||   |||||  |||||
Db       1202 CAGCTGGAGGAGGCTGAGGAAGGGGAGG 1229
```

```
RESULT 11
AB020967
LOCUS       AB020967               2004 bp    mRNA    linear   ROD 14-MAY-1999
DEFINITION  Rattus sp. mRNA for kinase, complete cds.
ACCESSION   AB020967
VERSION     AB020967.1  GI:4827158
KEYWORDS    kinase; NIPK.
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Mayumi-Matsuda,K., Kojima,S., Suzuki,H. and Sakata,T.
  TITLE     Identification of a novel kinase-like gene induced during neuronal
            cell death
  JOURNAL   Biochem. Biophys. Res. Commun. 258 (2), 260-264 (1999)
  MEDLINE   99262087
   PUBMED   10329375
REFERENCE   2  (bases 1 to 2004)
  AUTHORS   Matsuda,K.M., Kojima,S. and Nakayama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1998) Keiko Mayumi Matsuda, SHIONOGI and Co.
            Ltd., Shionogi Institute for Medical Science; 2-5-1 Mishima,
            Settsu-shi, Osaka 566-0022, Japan
            (E-mail:keiko.matsuda@shionogi.co.jp, Tel:81-6-6382-2612(ex.465),
            Fax:81-6-6382-2598)
FEATURES             Location/Qualifiers
     source          1. .2004
                     /organism="Rattus sp."
                     /mol_type="mRNA"
                     /db_xref="taxon:10118"
                     /cell_line="PC12"
                     /cell_type="pheochromocytoma"
     gene            1. .2004
                     /gene="NIPK"
     CDS             154. .1203
                     /gene="NIPK"
                     /note="neuronal cell death inducible putative kinase
                     (NIPK);induced by NGP-depletion"
                     /codon_start=1
                     /product="kinase"
                     /protein_id="BAA77582.1"
                     /db_xref="GI:4827159"
                     /translation="MRATSLAASADVPCRKKPLEPDDNIDVECPVLKRVRDEPEPGPT
                     PSLPPASDLSPAVAPATRLGPYILLEREQGNCTYRALHCPTGTEYTCKVYPASEAQAV
                     LAPYARLPTHQHVARPTEVLLGSQLLYTFPTKTHGDLHSLVRSRRGIPEPEAAALFRQ
                     MASAVAHCHKHGLILRDLKLRRFVFSNCERTKLVLENLEDACVMTGPDDSLWDKHACP
                     AYVGPEILSSRPSYSGRAADVWSLGVALFTMLAGRYPFQDSEPALLFGKIRRGTFALP
                     EGLSASARCLIRCLLRREPSERLVALGILLHPWLREDCSQVSPPRSDRREMDQVVPDG
                     PQLEEAEEGEVGLYG"
     polyA_site      2004
                     /gene="NIPK"
                     /note="27 a nucleotides"
BASE COUNT      410 a    563 c    590 g    441 t
ORIGIN
```

```
  Query Match           56.9%;  Score 611.2;  DB 10;  Length 2004;
  Best Local Similarity   74.9%;  Pred. No. 4.4e-103;
  Matches  785;  Conservative    0;  Mismatches  248;  Indels   15;  Gaps    1;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              ||||||||||  ||||||||||  |||| | |   |  | |||||||| | |||||||
Db        154 ATGCGAGCCACATCTCTGGCTGCTTCTGCCGATGTTCCCTGCAGGAAGAAACCGTTGGAG 213

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              || |||||||||  | |||  |||| |||| |||  |||||||   ||  ||  |||  ||
Db        214 TTTGATGACAACATTGATGTCGAGTGTCCAGTCTTAAAACGAGTGAGAGATGAGCCTGAG 273

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||  |||   |||||                  |||| |||| | ||  ||||    || ||
Db        274 CCTGGACCAACCCCC--------------AGCCTACCTCCCGCCTCAGACTTGTCACCT 318

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              |||||||| ||||  |||| ||  |||||||| ||||  |||| |  |||  | |||
Db        319 GCTGTGGCCCCTGCAACCCGGCTGGGGCCCTATATCCTTTTGGAACGAGAGCAAGGCAAC 378

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |  |||| ||||||||||||||||  ||||||||  |||||  |||||||||||||||  |||
Db        379 TGCACCTATCGGGCCCTGCACTGCCCCACAGGCACAGAGTACACCTGCAAGGTGTATCCC 438

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              | |   || |||| |||  |||||||  ||  ||  ||  | |||||  | |||  |||| |||
Db        439 GCCAGCGAGGCCCAGGCGGTGCTGGCACCGTATGCAAGACTGCCTACCCACCAGCACGTG 498

Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              || ||  |||||  |||||||||   ||  |  ||||||||  |||  ||||||||||   ||||
Db        499 GCCCGTCCCACAGAGGTCCTGCTGGGCTCTCAGCTCCTATACACCTTTTTCACGAAGACC 558

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||  |||||||  ||||||||  |||||||  |  |||||||  |||||  |||||||||
Db        559 CATGGGGACTTGCACAGTCTGGTGCGCAGCCGCCGCGGTATCCCCGAGCCCGAGGCTGCC 618

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              | |||  |||||  ||||||||  |  ||  ||||  |||||  |||  |||||||  ||   ||  ||
Db        619 GCGCTTTTCCGACAGATGGCTAGTGCGGTGGCACACTGCCACAAGCACGGGCTCATCTTG 678

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              ||  ||  ||||||  ||  ||||  |||||||||     ||  |||||||||  ||||  ||||||||
Db        679 CGCGACCTCAAGTTGCGTCGATTTGTCTTCAGCAACTGTGAGAGGACGAAGTTGGTGCTG 738

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              ||||||||||||||  ||||  |||  |||||  |||||  ||  ||  ||||||||||||
Db        739 GAGAACCTGGAGGATGCCTGTGTGATGACTGGACCAGACGACTCTCTGTGGGACAAGCAC 798

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              ||||||||  |||||  ||||||||  ||||||||||||||  |||  |  ||  |||||  ||||
Db        799 GCGTGCCCTGCCTATGTGGGACCAGAGATACTCAGCTCCCGGCCATCCTACTCCGGCAGA 858

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780
              ||  ||  ||||||||||||||||||||||||||||||||||||||||||||  ||||  |||||
Db        859 GCGGCTGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCTGGCCGATACCCG 918
```

```
Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
              |||||||||||  ||  ||  |  |||||||| ||||||||||   |  |||  |||  ||
Db        919 TTCCAGGACTCTGAACCAGCCCTGCTCTTTGGCAAGATCCGTAGAGGGACCTTTGCCCTG 978

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||  |||||  ||  |||  |||| ||  ||  ||  |  |||||  |||||  ||  |||||||
Db        979 CCTGAGGGCCTATCAGCCTCTGCTCGATGCCTCATCCGCTGTCTCCTCCGCAGGGAGCCT 1038

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              |  ||  ||  ||     |||    ||  |||||  ||  ||  ||||||  ||  |  ||||     |
Db       1039 TCAGAGCGACTTGTGGCCCTGGGAATCCTGCTACATCCCTGGTTGAGGGAGGATTGCAGC 1098

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              |    |  ||  |  ||  ||  |  |     |||||    |  ||||||  |||||  |||||  |
Db       1099 CAAGTCTCTCCTCCACGGTCTGACCGAAGGGAGATGGACCAGGTTGTCCCAGATGGGCCA 1158

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGG 1048
                |||||| ||  ||   |||||  |||||
Db       1159 CAGCTGGAGGAGGCAGAGGAAGGGGAGG 1186


RESULT 12
AX364906
LOCUS       AX364906               1076 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 57 from Patent WO0206315.
ACCESSION   AX364906
VERSION     AX364906.1  GI:18696795
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mintz,L., Freilich,S. and Bernstein,J.
  TITLE     Novel nucleic acid and amino acid sequences
  JOURNAL   Patent: WO 0206315-A 57 24-JAN-2002;
            Compugen Ltd. (IL)
FEATURES             Location/Qualifiers
     source          1. .1076
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      172 a    352 c    346 g    194 t     12 others
ORIGIN

  Query Match           55.6%;  Score 597.2;  DB 6;  Length 1076;
  Best Local Similarity   98.1%;  Pred. No. 1.9e-100;
  Matches  622;  Conservative    3;  Mismatches    6;  Indels    3;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 388

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 448

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 508

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 568

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 628

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 688
```

```
Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 748

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||||||||||||||||||||||||||::|||||||||||||||||||||||||||||
Db        749 CATGGGGACATGCACAGCCTGGTGCGAAGCSSCCACCGTATCCCTGAGCCTGAGGCTGCC 808

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 868

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        869 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG-AGAAGCTGGTGCTG 927

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 634
              ||||  |  |  |||||:|||||||||||||||
Db        928 GAGA--CCTGGAGACTCSTGCGTGCTGACTGGGC 959


RESULT 13
AX364921
LOCUS       AX364921               1076 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 72 from Patent WO0206315.
ACCESSION   AX364921
VERSION     AX364921.1  GI:18696810
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mintz,L., Freilich,S. and Bernstein,J.
  TITLE     Novel nucleic acid and amino acid sequences
  JOURNAL   Patent: WO 0206315-A 72 24-JAN-2002;
            Compugen Ltd. (IL)
FEATURES             Location/Qualifiers
     source          1. .1076
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      172 a    352 c    346 g    194 t     12 others
ORIGIN

  Query Match           55.6%;  Score 597.2;  DB 6;  Length 1076;
  Best Local Similarity   98.1%;  Pred. No. 1.9e-100;
  Matches  622;  Conservative    3;  Mismatches    6;  Indels    3;  Gaps    2;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 388

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 448

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 508

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 568

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 628

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 688
```

```
Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 748

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              ||||||||||||||||||||||||||||||::||||||||||||||||||||||||||||
Db        749 CATGGGGACATGCACAGCCTGGTGCGAAGCSSCCACCGTATCCCTGAGCCTGAGGCTGCC 808

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 868

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        869 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG-AGAAGCTGGTGCTG 927

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGC 634
              ||||  |  |  |||||:||||||||||||||||
Db        928 GAGA--CCTGGAGACTCSTGCGTGCTGACTGGGC 959
```

RESULT 14
BC012955

```
LOCUS       BC012955                1969 bp    mRNA    linear   ROD 16-APR-2003
DEFINITION  Mus musculus induced in fatty liver dystrophy 2, mRNA (cDNA clone
            MGC:18731 IMAGE:3980838), complete cds.
ACCESSION   BC012955
VERSION     BC012955.1  GI:15277944
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1969)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1969)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 23 Row: k Column: 9
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .1969
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:18731 IMAGE:3980838"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1969
                     /gene="Ifld2"
                     /note="synonym: TRB-3"
                     /db_xref="LocusID:228775"
                     /db_xref="MGI:1345675"
     CDS             177. .1163
                     /codon_start=1
                     /product="TRB-3"
                     /protein_id="AAH12955.1"
                     /db_xref="GI:15277945"
                     /db_xref="LocusID:228775"
                     /translation="MRATPLAASADVSCRKKPLEFDDNIDAKCPVLKRVRDEPEPGPL
                     PSLLPPSPPPASDLSPAVAPATRLGPYILLERRQGSCSYRALHCPTGTEYTCKVYPAS
                     EAQAVLAPYARLPTHQHVARPTEVLLGSRLLYIFPTKTHGDLHSLVRSRRGIPEPEAA
                     GLFRQMASAVAHCHKHGLVLRDLKLRRFVFSNCERTKLVLENLEDACVMTGSDDSLWD
                     KHACPAYVGPEILSSRPSYSEPVLLFGKIRRGTFALPEGLSAPARCLIRCLLRKEPSE
                     RLVALGILLHPWLREDHGRVSPPQSDRREMDQVVPDGPQLEEAEEGEVGLYG"
BASE COUNT      413 a    560 c    572 g    424 t
ORIGIN
```

```
  Query Match             47.7%;  Score 512.4;  DB 10;  Length 1969;
  Best Local Similarity   70.5%;  Pred. No. 7.8e-85;
  Matches  739;  Conservative    0;  Mismatches  231;  Indels   78;  Gaps    1;

Qy          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60
              |||||||| || |||||||||||| |||| | |      | |||||||| | |||||||
Db        177 ATGCGAGCTACACCTCTGGCTGCTTCTGCTGATGTTTCCTGCAGGAAGAAACCGTTGGAG 236

Qy         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120
              || ||||||||  | ||| || || |||| |||| |||||||   || || ||| ||
Db        237 TTTGATGACAATATTGATGCCAAGTGTCCAGTCCTAAAACGAGTGAGAGATGAGCCTGAG 296

Qy        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180
              ||| |||  | |||| ||||| ||||||  |||||||| | ||  ||||     || ||
Db        297 CCCGGACCACTCCCCAGCCTGCTGCCCCCCAGCCCACCTCCCGCCTCAGACTTGTCACCT 356

Qy        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
              |||||||| |||| | || || ||||| ||| |||| |||| | ||| | ||| |
Db        357 GCTGTGGCCCCTGCAACTCGACTGGGGCCTTATATCCTTTTGGAACGAGAGCAAGGCAGC 416

Qy        241 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300
              |   ||| || |||||||||||||| |||||||| ||||| ||||||||||||||||||
Db        417 TGCAGCTATCGAGCCCTGCACTGCCCCACAGGCACAGAGTACACCTGCAAGGTGTACCCT 476

Qy        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 360
              | |   || |||| ||| |||||||  || || || ||||||||  | ||| |||||||||
Db        477 GCCAGCGAGGCCCAGGCGGTGCTGGCACCTTATGCCCGGCTGCCTACCCACCAGCATGTG 536
```

```
Qy        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420
              || || ||||| |||||||||   ||  | | |||||| |||  |||||||||   ||||
Db        537 GCCCGTCCCACAGAGGTCCTGCTGGGCTCTCGGCTCCTTTACATCTTTTTCACGAAGACC 596

Qy        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480
              |||||||| |||||||||||||||| ||||||| | ||||||| ||||| |||||||||
Db        597 CATGGGGACTTGCACAGCCTGGTGCGCAGCCGCCGCGGTATCCCAGAGCCCGAGGCTGCC 656

Qy        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540
              | ||||||||| |||||||| |  ||| |||| ||||| ||| ||||||| || ||| ||
Db        657 GGGCTCTTCCGGCAGATGGCTAGTGCCGTGGCACACTGCCACAAGCACGGGCTTGTCTTG 716

Qy        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600
              || || ||||||||| ||||||||||||||    || ||||||||| ||||||||||||||
Db        717 CGCGACCTCAAGCTGCGTCGCTTTGTCTTCAGCAACTGTGAGAGGACGAAGCTGGTGCTG 776

Qy        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660
              |||||||||| ||  |||||||| |||||||  |||||| || ||||||||||||||||
Db        777 GAGAACCTGGAAGATGCCTGCGTGATGACTGGATCAGATGACTCTCTGTGGGACAAGCAT 836

Qy        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720
              |||||||| |||||||||||||| |||||||||||||| ||| | || |
Db        837 GCGTGCCCTGCCTACGTGGGACCAGAGATACTCAGCTCCCGGCCATCCT---------- 885

Qy        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Db        886 ------------------------------------------------------------ 885

Qy        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840
                  |||| ||||| |||||||||||| |||||||||  | ||| |||  ||| ||
Db        886 -------ACTCTGAGCCAGTCCTGCTCTTTGGCAAGATCCGTAGAGGGACCTTTGCCCTG 938

Qy        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900
              ||||  ||||| || ||||| |||||||||||| | ||||| ||||| ||   ||| ||
Db        939 CCTGAGGGCCTATCAGCCCCAGCCCGCTGTCTGATCCGCTGTCTCCTCCGCAAGGAACCT 998

Qy        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960
              | || || ||    |||   || |||||| |||| |||||| || || |||| |
Db        999 TCAGAGCGACTTGTGGCCCTGGGAATCCTCTTGCATCCCTGGTTGAGAGAGGATCACGGC 1058

Qy        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020
              |   | || |  | |  ||  ||   |||||   | ||||||||||| ||||| |
Db       1059 CGAGTCTCTCCTCCACAGTCTGACCGAAGGGAGATGGACCAGGTGGTCCCAGATGGGCCA 1118

Qy       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGG 1048
                ||||||  || ||   ||||| |||||
Db       1119 CAGCTGGAGGAGGCTGAGGAAGGGGAGG 1146
```

RESULT 15
HS1103G7

LOCUS	HS1103G7 153170 bp DNA linear PRI 08-FEB-2001
DEFINITION	Human DNA sequence from clone RP5-1103G7 on chromosome 20p12.2-13. Contains up to three novel genes, the gene for a novel protein similar to mouse VMP, the gene for a novel protein kinase domains containing protein similar to phosphoprotein C8FW and rat NIPK, and the SOX22 gene for SRY (sex-determining region Y)-box 22. Contains five CpG islands, ESTs, STSs and GSSs, complete sequence.
ACCESSION	AL034548
VERSION	AL034548.25 GI:7263904
KEYWORDS	HTG; CpG island; NIPK; protein kinase; SOX22; SRY; VMP.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 153170)
AUTHORS	Blakey,S.
TITLE	Direct Submission
JOURNAL	Submitted (05-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT	On Mar 19, 2000 this sequence version replaced gi:5541861. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence is the entire insert of clone RP5-1103G7 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr20 RP5-1103G7 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pCYPAC2.

FEATURES	Location/Qualifiers
source	1. .153170 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="20" /map="p12.2-13" /clone="RP5-1103G7" /clone_lib="RPCI-5"
repeat_region	1. .61 /note="Alu repeat: matches 2. .62 of consensus"
repeat_region	79. .129 /note="L1MA9 repeat: matches 6221. .6271 of consensus"
repeat_region	130. .173 /note="22 copies 2 mer ag 79% conserved"
repeat_region	175. .873 /note="L1MA9 repeat: matches 5519. .6217 of consensus"
repeat_region	902. .1109 /note="L1MC5 repeat: matches 7715. .7921 of consensus"
repeat_region	1110. .1480 /note="THE1C repeat: matches 1. .371 of consensus"
misc_feature	1502. .1733 /note="match: STS: Em:Z66755"
repeat_region	1734. .1888 /note="MLT1H repeat: matches 387. .547 of consensus"
repeat_region	1919. .2101 /note="MLT1H repeat: matches 115. .298 of consensus"
misc_feature	complement(join(2146. .2243,126791. .126834)) /note="match: GSS: Em:B59651"
repeat_region	2154. .2193 /note="20 copies 2 mer aa 85% conserved"
misc_feature	complement(join(2194. .2243,126792. .126871)) /note="match: GSS: Em:AQ377604"
repeat_region	2699. .2917 /note="L1M4 repeat: matches 2875. .3081 of consensus"
repeat_region	2938. .3062 /note="L1M4 repeat: matches 3120. .3249 of consensus"
repeat_region	3136. .3430 /note="AluSg repeat: matches 1. .295 of consensus"
repeat_region	3482. .3777 /note="AluSq repeat: matches 1. .296 of consensus"
repeat_region	3811. .4126 /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region	4310. .4608 /note="L1MD2 repeat: matches 5861. .6167 of consensus"
repeat_region	4609. .5026 /note="MSTA repeat: matches 1. .424 of consensus"
repeat_region	5027. .5200

```
                     /note="L1MD2 repeat: matches 5684. .5861 of consensus"
repeat_region        5201. .5501
                     /note="AluSx repeat: matches 1. .307 of consensus"
repeat_region        5502. .5727
                     /note="L1MD2 repeat: matches 5451. .5684 of consensus"
repeat_region        5798. .5865
                     /note="34 copies 2 mer aa 66% conserved"
repeat_region        6056. .6643
                     /note="MER67C repeat: matches 122. .710 of consensus"
repeat_region        6648. .6707
                     /note="30 copies 2 mer ca 88% conserved"
repeat_region        6710. .6824
                     /note="MER50 repeat: matches 10. .134 of consensus"
repeat_region        6825. .7124
                     /note="AluSx repeat: matches 7. .306 of consensus"
repeat_region        7126. .7415
                     /note="L1MA4A repeat: matches 5619. .5919 of consensus"
repeat_region        7419. .7569
                     /note="AluSg/x repeat: matches 148. .298 of consensus"
misc_feature         7576. .7889
                     /note="match: GSS: Em:AQ180303"
misc_feature         complement(7590. .7894)
                     /note="match: GSS: Em:AQ701808"
misc_feature         7862. .8140
                     /note="match: GSS: Em:AQ123077"
misc_feature         8019. .8317
                     /note="match: GSS: Em:AQ746749"
misc_feature         8019. .8151
                     /note="match: GSS: Em:AQ617736"
repeat_region        8152. .8316
                     /note="MER4B repeat: matches 416. .574 of consensus"
repeat_region        8349. .8643
                     /note="MER4D repeat: matches 362. .689 of consensus"
repeat_region        8644. .8942
                     /note="AluY repeat: matches 1. .299 of consensus"
repeat_region        8943. .9113
                     /note="MER4D repeat: matches 194. .362 of consensus"
repeat_region        9113. .9201
                     /note="MLT2FA repeat: matches 377. .450 of consensus"
repeat_region        9226. .9311
                     /note="L1MB4 repeat: matches 5775. .5862 of consensus"
repeat_region        9312. .9651
                     /note="MER1B repeat: matches 1. .337 of consensus"
repeat_region        9653. .9750
                     /note="L1MA5 repeat: matches 5856. .5947 of consensus"
repeat_region        9751. .10047
                     /note="AluSx repeat: matches 1. .298 of consensus"
repeat_region        10048. .10369
                     /note="L1MA5 repeat: matches 5947. .6300 of consensus"
repeat_region        10488. .10529
                     /note="21 copies 2 mer at 88% conserved"
repeat_region        10532. .10653
                     /note="L1MC4 repeat: matches 7673. .7800 of consensus"
repeat_region        10687. .10752
                     /note="33 copies 2 mer at 68% conserved"
repeat_region        11681. .11742
                     /note="31 copies 2 mer ta 71% conserved"
misc_feature         13519. .13972
                     /note="match: GSS: Em:AQ224309"
repeat_region        14208. .14554
                     /note="LTR16B repeat: matches 110. .462 of consensus"
repeat_region        15554. .15872
                     /note="L1ME1 repeat: matches 5615. .5910 of consensus"
repeat_region        15873. .16174
                     /note="AluSp repeat: matches 16. .313 of consensus"
repeat_region        16175. .16382
                     /note="L1ME1 repeat: matches 5400. .5615 of consensus"
repeat_region        16383. .16423
                     /note="L1ME repeat: matches 5893. .5933 of consensus"
repeat_region        16595. .16894
                     /note="AluY repeat: matches 1. .302 of consensus"
repeat_region        16895. .17021
                     /note="L1MA10 repeat: matches 6196. .6318 of consensus"
repeat_region        17155. .17587
                     /note="L1ME repeat: matches 5289. .5733 of consensus"
repeat_region        17587. .18157
                     /note="L1M1 repeat: matches -1389. .-790 of consensus"
repeat_region        18158. .18195
                     /note="19 copies 2 mer tg 97% conserved"
repeat_region        18197. .18230
                     /note="AluSg/x repeat: matches 165. .197 of consensus"
repeat_region        18231. .18823
                     /note="MER52A repeat: matches 1. .672 of consensus"
misc_feature         complement(18827. .19180)
                     /note="match: GSS: Em:AQ599398"
repeat_region        19689. .19998
                     /note="AluSg repeat: matches 1. .309 of consensus"
repeat_region        20287. .20543
                     /note="AluSx repeat: matches 2. .307 of consensus"
repeat_region        20552. .20609
                     /note="MLT1J repeat: matches 255. .318 of consensus"
repeat_region        20592. .20756
                     /note="MLT1E repeat: matches 256. .421 of consensus"
repeat_region        20895. .21012
                     /note="L1PBa repeat: matches -1128. .-1008 of consensus"
repeat_region        21009. .21168
                     /note="L1PBa repeat: matches -1537. .-1377 of consensus"
repeat_region        21192. .21277
                     /note="L1MA7 repeat: matches 5864. .5949 of consensus"
repeat_region        21280. .21627
                     /note="L2 repeat: matches 2250. .2617 of consensus"
misc_feature         21628. .21884
                     /note="match: GSS: Em:B41180"
misc_feature         21628. .21744
                     /note="match: GSS: Em:AQ058160"

Query Match             46.1%;  Score 495;  DB 9;  Length 153170;
Best Local Similarity   97.1%;  Pred. No. 5.5e-82;
Matches  504;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy        556 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 615
              || | |||    ||| |    ||  ||||||||||||||||||||||||||||||||
Db     150847 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 150906

Qy        616 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     150907 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 150966

Qy        676 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     150967 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 151026

Qy        736 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151027 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 151086

Qy        796 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151087 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 151146

Qy        856 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151147 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 151206

Qy        916 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151207 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 151266

Qy        976 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151267 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 151326

Qy       1036 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1074
              |||||||||||||||||||||||||||||||||||||||
Db     151327 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 151365
```

Search completed: January 16, 2004, 14:49:08
Job time : 4032.84 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 10:03:56 ; Search time 4616.16 Seconds
(without alignments)
10840.821 Million cell updates/sec

Title: US-09-909-474D-1
Perfect score: 2059
Sequence: 1 gctctgagccccggcggcgc.....aaaaaaaaaaaaaaaaaaaa 2059

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	940.6	45.7	1201	9	AL578892	AL578892 AL578892
c	2	902	43.8	942	13	BX393466	BX393466 BX393466
c	3	895.2	43.5	1075	13	BX367264	BX367264 BX367264
c	4	893	43.4	1051	9	AL525890	AL525890 AL525890
c	5	892	43.3	947	13	BX325344	BX325344 BX325344
	6	868.8	42.2	975	10	BG575275	BG575275 602597925
c	7	862.4	41.9	1001	9	AL562576	AL562576 AL562576
	8	849.4	41.3	873	13	BQ216198	BQ216198 AGENCOURT
	9	846.2	41.1	881	13	BQ941789	BQ941789 AGENCOURT
	10	839.8	40.8	945	13	BX325345	BX325345 BX325345
c	11	834	40.5	1201	13	BX426485	BX426485 BX426485
c	12	832.8	40.4	1201	13	BX445733	BX445733 BX445733
	13	818.6	39.8	863	10	BG748142	BG748142 602705453
	14	818.4	39.7	1201	9	AL556690	AL556690 AL556690
	15	816.6	39.7	948	13	BX443303	BX443303 BX443303
	16	814.6	39.6	902	13	BX421477	BX421477 BX421477
	17	811.2	39.4	1062	12	BM557568	BM557568 AGENCOURT
	18	807.2	39.2	837	12	BI861766	BI861766 603391172
	19	806.2	39.2	918	13	BQ223497	BQ223497 AGENCOURT
	20	802.4	39.0	1036	13	BQ217691	BQ217691 AGENCOURT
c	21	797.8	38.7	925	13	BX435881	BX435881 BX435881
	22	796.6	38.7	946	13	BX390085	BX390085 BX390085
	23	790.2	38.4	929	13	BU500619	BU500619 AGENCOURT
	24	788	38.3	929	9	AL522632	AL522632 AL522632
	25	765.6	37.2	1126	13	BX405957	BX405957 BX405957
	26	759	36.9	944	13	BX363066	BX363066 BX363066
	27	756	36.7	914	13	BU500176	BU500176 AGENCOURT
	28	753.8	36.6	1076	13	BX367265	BX367265 BX367265
	29	744.6	36.2	851	12	BI909168	BI909168 603069177
	30	744	36.1	820	12	BI256421	BI256421 602974439
	31	740.6	36.0	880	10	BG470123	BG470123 602533821
	32	738.6	35.9	1002	12	BM804790	BM804790 AGENCOURT
c	33	737	35.8	910	9	AL562114	AL562114 AL562114
	34	736.4	35.8	892	14	CA495347	CA495347 AGENCOURT
	35	735.8	35.7	1053	13	BQ224844	BQ224844 AGENCOURT
c	36	735.2	35.7	861	9	AL555266	AL555266 AL555266
	37	735	35.7	989	10	BE792929	BE792929 601585274
c	38	730.2	35.5	1051	12	BM907602	BM907602 AGENCOURT
	39	728.2	35.4	810	12	BM045379	BM045379 603623380
	40	725.4	35.2	838	10	BG171687	BG171687 602322210
	41	718	34.9	792	10	BG387820	BG387820 602412721
	42	714.6	34.7	931	10	BG470193	BG470193 602533702
	43	713.2	34.6	957	13	BQ949069	BQ949069 AGENCOURT
c	44	712.8	34.6	918	9	AL520575	AL520575 AL520575
	45	709.4	34.5	910	10	BG397691	BG397691 602438746

ALIGNMENTS

RESULT 1
AL578892/c
LOCUS AL578892 1201 bp mRNA linear EST 01-JUN-2003
DEFINITION AL578892 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens cDNA clone CS0DK005YC05 3-PRIME, mRNA sequence.
ACCESSION AL578892
VERSION AL578892.2 GI:31317080
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced gi:12943405.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10000.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK005AB03NP1&cluster=10000.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :

```
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK005AB03NP1.
FEATURES              Location/Qualifiers
     source           1. .1201
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DK005YC05"
                      /cell_type="HELA CELLS COT 25-NORMALIZED"
                      /cell_line="HELA"
                      /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      285 a    298 c    288 g    246 t     84 others
ORIGIN

  Query Match             45.7%;  Score 940.6;  DB 9;  Length 1201;
  Best Local Similarity   96.1%;  Pred. No. 4.8e-77;
  Matches  943;  Conservative   24;  Mismatches   13;  Indels    1;  Gaps    1;

Qy        984 CCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGC 1043
              :||||:|:  ||||:: |||||:|::||||||:: ||    ||| :||||:|||| :||||||
Db        992 MCCCTKGSCGCGAMWAGACCCKAKSCCCTTDSTCCACCCCCTWTCCCWTCTC-KGGAGGC 934

Qy       1044 TGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAG 1103
              ||||:|||||  |||  |: |:|: |:||||||||||||||||||||||||||||||||||
Db        933 TGCCYAGGTGTTCCTTKTTKGWTTKGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAG 874

Qy       1104 AGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAGCTGCCAACAGTGGATT 1163
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db         873 AGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAKCTGCCAACAGTGGATT 814

Qy       1164 GAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCC 1223
              ||||||||||||||||||||||||||||||:|||:||||:||||||||||||||||||||
Db         813 GAGTTTGGGGGTAGCTCCAAGCCTTCTCYTGCYTCTGWACTGAGCCAAACCTTCAGTGCC 754

Qy       1224 TTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGT 1283
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db         753 TTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGTACAYATCTGCTTTGT 694

Qy       1284 TCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGC 1343
              ||||||||||||||||||||||||||||||||||||||||||||:|||||| |||||
Db         693 TCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGCCYTGTTCTAGGTGC 634

Qy       1344 TGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGCTCACAGAGATGACAAACTGG 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         633 TGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGCTCACAGAGATGACAAACTGG 574

Qy       1404 CATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCACTGTCTACACTGGGTACACT 1463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         573 CATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCACTGTCTACACTGGGTACACT 514

Qy       1464 TTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATC 1523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         513 TTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATC 454

Qy       1524 CCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTAT 1583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         453 CCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTAT 394

Qy       1584 CCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTC 1643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         393 CCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTC 334

Qy       1644 TGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTA 1703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         333 TGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTA 274

Qy       1704 AGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAGTTCTTGTCTAACTC 1763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         273 AGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAGTTCTTGTCTAACTC 214

Qy       1764 AAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGGCTTCTGACCTGAGCACC 1823
              ||||:| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         213 AAGACTATTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGGCTTCTGACCTGAGCACC 154

Qy       1824 AAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTG 1883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         153 AAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTG 94

Qy       1884 GGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACTCTAGGTTTTGGATA 1943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          93 GGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACTCTAGGTTTTGGATA 34

Qy       1944 CCATGAGTATGTATGTTTACC 1964
              |||||||||||||||||||||
Db          33 CCATGAGTATGTATGTTTACC 13


RESULT 2
BX393466/c
LOCUS       BX393466                942 bp    mRNA    linear   EST 13-MAY-2003
DEFINITION  BX393466 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
            cDNA clone CS0DB003YP19 3-PRIME, mRNA sequence.
ACCESSION   BX393466
VERSION     BX393466.1  GI:30620023
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 942)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DB003CH10NP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DB003CH10NP1.
FEATURES              Location/Qualifiers
     source           1. .942
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DB003YP19"
                      /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
                      /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      224 a    248 c    243 g    210 t     17 others
ORIGIN

  Query Match             43.8%;  Score 902;  DB 13;  Length 942;
  Best Local Similarity   97.0%;  Pred. No. 1.7e-73;
  Matches  906;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1035 CTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGAAGAGGA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         942 CTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGAAGAGGA 883

Qy       1095 GGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAGCTGCCAA 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         882 GGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAGCTGCCAA 823
```

```
Qy       1155 CAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGCCAAACC 1214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGCCAAACC 763

Qy       1215 TTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGTACACAT 1274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         762 TTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGTACACAT 703

Qy       1275 CTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGCCCTGT 1334
             |||||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db         702 CTGCTTTGTTCCACACACATTSAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGCCCTGT 643

Qy       1335 TCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGCTCACAGAGATG 1394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         642 TCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGCTCACAGAGATG 583

Qy       1395 ACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCACTGTCTACACT 1454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         582 ACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCACTGTCTACACT 523

Qy       1455 GGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCA 1514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         522 GGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCA 463

Qy       1515 AGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAGAGAAA 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         462 AGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAGAGAAA 403

Qy       1575 GGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGACCCAAGCCC 1634
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         402 GGGAGGTATCCCTGTGCCAAAGGCTCTAGGCCTCTCCCCTGCAACTCAGGACCCAAGCCC 343

Qy       1635 AGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCTT 1694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         342 AGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCTT 283

Qy       1695 CCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAGTTCTT 1754
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         282 CCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAACCCAAACTATGAGGCTAGTTCTT 223

Qy       1755 GTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGGCTTCTGAC 1814
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db         222 GTCTAACTCAAGACTGTTCTGGAATGAGGGWCCAGGCCTGTCAACCATGGGGCTTCTGAC 163

Qy       1815 CTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGAAAG 1874
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         162 CTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGAAAG 103

Qy       1875 TCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACTCTAGG 1934
             |||||||||||||| |||||||:|| ||||| |:|||||||||| ||||||||||||
Db         102 CCCCAGGTGGGACTCGTCTGGGGASACGTGGGGNCSACAATCCCAGGGCCATACTCTAGG 43

Qy       1935 TTTTGGATACCATGAGTATGTATGTTTACCTGTG 1968
             :  |   | ||:   || ||| ||| ||||| |:::
Db         42 YCCTNNANACSGGGATTATTTATTTTTACTTKKB 9


RESULT 3
BX367264/c
LOCUS       BX367264               1075 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  BX367264 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
            Homo sapiens cDNA clone CS0DL002YJ07 3-PRIME, mRNA sequence.
ACCESSION   BX367264
VERSION     BX367264.1  GI:30441034
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1075)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AL002CE04NP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AL002CE04NP1.
FEATURES             Location/Qualifiers
     source          1. .1075
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DL002YJ07"
                     /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                     25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      264 a    256 c    281 g    243 t     31 others
ORIGIN

  Query Match           43.5%;  Score 895.2;  DB 13;  Length 1075;
  Best Local Similarity   95.7%;  Pred. No. 6.7e-73;
  Matches  929;  Conservative   10;  Mismatches   30;  Indels    2;  Gaps    2;

Qy       1014 AGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCT 1073
             | |||  |: | | || : |  |     ||     ||      |  ||  |||||||||||| |
Db         970 ATCCCTTASTCAACCCAWCCCATTTTGGGAGCTCCCCAGTGGTCCCTATGGACTGGGGGT 911

Qy       1074 GGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCC 1133
             :||||||:|||||||||||||| |||:||||||||:|||||||:|||||||:||||||||
Db         910 KGACGAARCCAGGGAAGAGGA-GGARACAGAGAAKTGGTTCTKTATGGCTARGACCACCC 852

Qy       1134 TACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCT 1193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         851 TACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCT 792

Qy       1194 GCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTG 1253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         791 GCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTG 732

Qy       1254 TGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGC 1313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         731 TGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGC 672

Qy       1314 TTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAA 1373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         671 TTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAA 612

Qy       1374 TATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGA 1433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         611 TATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGA 552

Qy       1434 CCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTG 1493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         551 CCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTG 492

Qy       1494 GTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTA 1553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         491 GTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTA 432

Qy       1554 TCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCC 1613
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         431 TCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCTAGGCCTCTCCCC 372
```

```
Qy       1614 TGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCT 1673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 TGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCT 312

Qy       1674 CTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATC 1733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 CTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATC 252

Qy       1734 CAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCT 1793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCT 192

Qy       1794 GTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCT 1853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCT 132

Qy       1854 GTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACA 1913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 GTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACA 72

Qy       1914 ATCCCAGGTCCATACTCTAGGTTTTGGATACC-ATGAGTATGTATGTTTACCTGTGCCTA 1972
             |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||::||
Db          71 ATCCCAGGTCCATACTCTAGGTTTTGGATACCNATGAGTATGTATGTTTACCTNTGBYTA 12

Qy       1973 ATAAAGGAGAA 1983
             ||||||||||||
Db          11 ATAAAGGAGAA 1
```

```
RESULT 4
AL525890/c
LOCUS       AL525890                1051 bp    mRNA    linear   EST 23-MAY-2003
DEFINITION  AL525890 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC013YJ08 3-PRIME, mRNA sequence.
ACCESSION   AL525890
VERSION     AL525890.2  GI:31063754
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1051)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 13, 2001 this sequence version replaced gi:12789383.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DC013DE04NP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DC013DE04NP1.
FEATURES             Location/Qualifiers
     source          1. .1051
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC013YJ08"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      240 a    259 c    278 g    233 t     41 others
ORIGIN
```

```
  Query Match           43.4%;  Score 893;  DB 9;  Length 1051;
  Best Local Similarity  95.7%;  Pred. No. 1.1e-72;
  Matches  908;  Conservative   25;  Mismatches   14;  Indels    2;  Gaps    2;

Qy       1010 CCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGG 1069
             ||||||| |||:|||||||||||||||||||||||:||||||||::|||||||||||||||
Db        949 CCTTAGCTCCAMCCCGATCCCATCTCTGGGAGGCTSCCCAGGTGKYCCCTGATGGACTGG 890

Qy       1070 GGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACC 1129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACC 830

Qy       1130 ACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTC 1189
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 ACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTC 770

Qy       1190 TCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCC 1249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 TCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCC 710

Qy       1250 TGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGG 1309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 TGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGG 650

Qy       1310 GTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAG 1369
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        649 GTGCTTATCAGGTGCCAAGCCCTGTTCTAGGTGCTGGGAGTACAGCAGTGAGCAAAGGAG 590

Qy       1370 ACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCC 1429
             ||||||||||||||||:||||||:|:||||||||||||||||||||||||||||||||||||
Db        589 ACAATATTCCCTGCTCAMAGAGATRAMAAACTGGCATCCTTGAGCTGACAACACTTTTCC 530

Qy       1430 ATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGAT 1489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGAT 470

Qy       1490 GCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTT 1549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTT 410

Qy       1550 TGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCT 1609
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        409 TGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCACCAGGCCTCT 350

Qy       1610 CCCCTGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTG 1669
             |:|: :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 CMCMMSCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTG 290

Qy       1670 TCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAG 1729
             ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        289 TCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGVATTTGGGCCAGAGATAAG 230

Qy       1730 AATCCAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAG 1789
             ||:|||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        229 AAYCCAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTATTCTGGAATGAGGGTCCAG 170

Qy       1790 GCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGG 1849
             ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 GCCTBTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGG 110

Qy       1850 GTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTC 1909
             |||||||||||||||||||||||||||||||||||||||||||||||| ||:||:| || ||
Db        109 GTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGNGABACKT-GGNTC 51

Qy       1910 CACAATCCCAGGTCCATACTCTAGGTTTTGGATACCATGAGTATGTATG 1958
              :| |||:||:|||| ||| | ||||| || |||:| | |||:|:::|::
Db         50 BAAAATSCCVGGTCNATAGTNTAGGTATT-GATRCGAGGAGWAKRKAKR 3
```

```
RESULT 5
BX325344/c
LOCUS       BX325344                947 bp    mRNA    linear   EST 02-MAY-2003
DEFINITION  BX325344 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
```

```
            Homo sapiens cDNA clone CS0DL004YF04 3-PRIME, mRNA sequence.
ACCESSION   BX325344
VERSION     BX325344.1  GI:30336457
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 947)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AL004DC02NP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AL004DC02NP1.
FEATURES             Location/Qualifiers
     source          1. .947
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DL004YF04"
                     /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                     25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      232 a    241 c    241 g    205 t     28 others
ORIGIN

  Query Match           43.3%;  Score 892;  DB 13;  Length 947;
  Best Local Similarity  96.4%;  Pred. No. 1.4e-72;
  Matches  907;  Conservative   22;  Mismatches   10;  Indels    2;  Gaps    2;

Qy       1030 CATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGAA 1089
             | |||||||||||||: |||||::||||||||:||||||||||:||:||:||:||||||||
Db        947 CTTCTCTGGGAGGCKCCCCAGKKGGTCCCTGWTGGACTGGGGYTGRACRAASCCAGGGAA 888

Qy       1090 GAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAGCT 1149
             :|||||:|||:|||||||||||||||:||||||||||||||||||||||||||||||||||
Db        887 RAGGAGGRAGAMAGAGAAGTGGTTCTKTATGGCTAGGACCACCCTACTACACGCTCAGCT 828

Qy       1150 GCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGCC 1209
             ||||||||||||||||||||||||||:||||||||||||||:|||||||||||||||||||
Db        827 GCCAACAGTGGATTGAGTTTGGGGGTRGCTCCAAGCCTTCTCCTSCCTCTGAACTGAGCC 768

Qy       1210 AAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGTA 1269
             |||||||:|||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        767 AAACCTTHAGTGCCTTCCAGAAGGGAGAAAGGCARAAGCCTGTGTGGAGTGTGCTGTGTA 708

Qy       1270 CACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGC 1329
             ||||||||||||||:|||||||  :||||||||||||||||||||||||||||||||||||
Db        707 CACATCTGCTTTGTTCMACACACAATSAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGC 648

Qy       1330 CCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGCTCACAG 1389
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        647 CCTGTTCTCGGTGCTGGGAGWACAGCAGTGAGCAAAGGAGACAATATTCCCTGCTCACAG 588

Qy       1390 AGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCACTGTCT 1449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCACTGTCT 528

Qy       1450 ACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCACCTCT 1509
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        527 ACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCT-GTGCTCAGGCACCTCT 469

Qy       1510 GTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAG 1569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAG 409

Qy       1570 AGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGACCCA 1629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 AGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGACCCA 349

Qy       1630 AGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTC 1689
             |||||:|||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        348 AGCCCMGCTCACTCTGGGTACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTC 289

Qy       1690 TCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAG 1749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 TCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAG 229

Qy       1750 TTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGGCTT 1809
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 TTCTTGTCTTACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGGCTT 169

Qy       1810 CTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTG 1869
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTG 109

Qy       1870 GAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACT 1929
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 GAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATACT 49

Qy       1930 CTAGGTTTTGGATAC-CATGAGTATGTATGTTTACCTGTGC 1969
             ||||||||||||||| |||| |||||||||||||  :  |||:
Db          48 CTAGGTTTTGGATACNCATGGGTATGTATGTTTNBKGGTGS 8

RESULT 6
BG575275
LOCUS       BG575275                975 bp    mRNA    linear   EST 10-APR-2001
DEFINITION  602597925F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4706575 5',
            mRNA sequence.
ACCESSION   BG575275
VERSION     BG575275.1  GI:13582928
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 975)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10575  row: a  column: 08
            High quality sequence stop: 891.
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4706575"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_87"
```

```
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      199 a    301 c    262 g    213 t
ORIGIN

  Query Match             42.2%;  Score 868.8;  DB 10;  Length 975;
  Best Local Similarity   96.2%;  Pred. No. 1.7e-70;
  Matches  922;  Conservative    0;  Mismatches   32;  Indels    4;  Gaps    3;

Qy        713 GCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAG 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAG 60

Qy        773 CCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCC 832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCC 120

Qy        833 AGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTG 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTG 180

Qy        893 CAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTG 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTG 240

Qy        953 AACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCT 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCT 300

Qy       1013 TAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGC 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGC 360

Qy       1073 TGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACC 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACC 420

Qy       1133 CTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCC 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCC 480

Qy       1193 TGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGT 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGT 540

Qy       1253 GTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTG 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTG 600

Qy       1313 CTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACA 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACA 660

Qy       1373 ATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATG 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATG 720

Qy       1433 ACCATAGGTCACTGTCTACACTGGGTACACTTT--GTAC[illegible]GGCCTCCACTGATG 1490
              ||||||||||||||||||||||||||||||||||      |||||||| ||||||||||
Db        721 ACCATAGGTCACTGTCTACACTGGGTACACTTTTGTACCCAGTGTCGGTCTCCACTGATG 780

Qy       1491 CTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTT 1550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACAGCTTGCCTTT 840

Qy       1551 -GTATCTTGTACCTTTTC-AGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 1608
               ||||| |||||||||||| ||| || |||||| ||||||| | |||||||| |  | |||
Db        841 GGTATCCTGTACCTTTTCAAGATAAGGGGAGGAATCCCTGGGGCAAAGGCTTCCAGGCTC 900

Qy       1609 TCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCT 1666
              | |||    | ||| |||||||||||||  |||||||||||||| ||||||||   ||
Db        901 TTCCCCTGGAATCAAGACCCAAGCCCAGTCAACTCTGGGAACTGGGTTCCCAGATCCT 958


RESULT 7
AL562576/c
LOCUS       AL562576                1001 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL562576 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC013YJ21 3-PRIME, mRNA sequence.
ACCESSION   AL562576
VERSION     AL562576.2  GI:31286588
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12911134.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DC013CE11NP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DC013CE11NP1.
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC013YJ21"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      230 a    250 c    257 g    224 t     40 others
ORIGIN

  Query Match             41.9%;  Score 862.4;  DB 9;  Length 1001;
  Best Local Similarity   94.0%;  Pred. No. 6.5e-70;
  Matches  942;  Conservative   33;  Mismatches   18;  Indels    9;  Gaps    8;

Qy        969 AGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATC 1028
              | ||||||||||||||||:|||:||||:|||||| ||||||||||: ||||||||||
Db       1000 AAGCATCCTCCTGCACCCYTGGSTGCGAYAGGACCCTATGCCCTTAGYTCCAACCCGATC 941

Qy       1029 CCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGA 1088
              ::||||:|||||||||||||: || ||||||||||||||||||||||||||||||||||
Db        940 YYATCTYTGGGAGGCTGCCCHAGTTGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGA 881

Qy       1089 AGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAGC 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 AGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAGC 821

Qy       1149 TGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGC 1208
              |:|||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        820 TKCCAACAGTGGATTGAGTTT-GGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGC 762

Qy       1209 CAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGT 1268
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        761 CAAACCTTCAGTGCCTTCCAGAAGGGAGAAAKGCAGAAGCCTGTGTGGAGTGTGCTGTGT 702

Qy       1269 ACACATCTGCTTTGTTCCACACA-CATGCAGTTCCTGCTTGGGTGCTTATCAGGTGCCAA 1327
```

```
                |||||:|::|:::|::: ||||| |||||||||||||||:||||||||||||||||||
Db        701 ACACAYCYSCHHDGKKSGACACACCATGCAGTTCCTGCTTGKGTGCTTATCAGGTGCCAA 642

Qy       1328 GCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATAT-TCCCTGCTCA 1386
                |||:|||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        641 GCCYTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATATCCCTGCTCA 582

Qy       1387 CAGAGATGACAAACTG-GCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTC-AC 1444
                |||||||||||||||| |||||||||||||||||||||||||||||||||||||| ||
Db        581 CAGAGATGACAAACTGTGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCTAC 522

Qy       1445 TGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCA 1504
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 TGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCA 462

Qy       1505 CCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCC--TTTGTATCTTGTACC 1562
                |||||||||||||||||||||||||||||||||||||||||||  ||  ||||||||||
Db        461 CCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCATTTGATATCTTGTACC 402

Qy       1563 TTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCA 1622
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 TTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCTAGGCCTCTCCCCHGCAACTCA 342

Qy       1623 GGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAA 1682
                |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        341 GGACCCAAGCCCAGCTCACTCTGGGAACTGTGTBCCCAGCATCTCTGTCCTCTTGATTAA 282

Qy       1683 GAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATG 1742
                |||||||||||||||||||||||||||||:|||||||||||||||||:|||||||||||
Db        281 GAGATTCTCCTTCCAGGCCTAAGCCTGGRATTTGGGCCAGAGATAAGAAHCCAAACTATG 222

Qy       1743 AGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGT-CCAGGCCTGTCAACCA 1801
                |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        221 AGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTACCAGGCCTGTCAACCA 162

Qy       1802 TGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTG 1861
                |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        161 TGGGGCTTCTGACCTGAGCACCAAGKTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTG 102

Qy       1862 GCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGG 1921
                |||||||||||||||||||||||||||||||||||:||||||||:::|||||:|
Db        101 GCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGASACTTGGGGTCSMSAATCCCMGA 42

Qy       1922 TCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTAC 1963
                |  ||   |||||| | ||| ||||||||||||:||||||:
Db          41 CCNNTANNNTAGGTTNT-GATKCCATGAGTATGTMTGTTTAB 1
```

```
RESULT 8
BQ216198
LOCUS       BQ216198                873 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7574831 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059905
            5', mRNA sequence.
ACCESSION   BQ216198
VERSION     BQ216198.1  GI:20397598
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 873)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13327  row: f  column: 02
            High quality sequence stop: 640.
FEATURES             Location/Qualifiers
     source          1. .873
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6059905"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
BASE COUNT      144 a    305 c    264 g    160 t
ORIGIN
```

```
  Query Match             41.3%;  Score 849.4;  DB 13;  Length 873;
  Best Local Similarity   99.2%;  Pred. No. 1.1e-68;
  Matches  864;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 62

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 122

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 182

Qy        181 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 242

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        243 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 302

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 362

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 422

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 482

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 542

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 602

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 662

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 722

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 782
```

```
Qy        781 T-GGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTC 839
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 TGGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTC 842

Qy        840 GGAGCCTGTCCTGCTCTTCGGCAAGATCCGC 870
              |||||||||||||||||||| || |  | |
Db        843 GGAGCCTGTCCTGCTCTTCGGAAAAAAACCC 873


RESULT 9
BQ941789
LOCUS       BQ941789                881 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8821009 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6422656
            5', mRNA sequence.
ACCESSION   BQ941789
VERSION     BQ941789.1  GI:22357267
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 881)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2599  row: p  column: 17
            High quality sequence stop: 689.
FEATURES             Location/Qualifiers
     source          1. .881
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6422656"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_18"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT      161 a    275 c    262 g    181 t      2 others
ORIGIN

  Query Match             41.1%;  Score 846.2;  DB 13;  Length 881;
  Best Local Similarity   98.6%;  Pred. No. 2.1e-68;
  Matches  863;  Conservative    0;  Mismatches   10;  Indels    2;  Gaps    1;

Qy        531 GCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCG 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCG 60

Qy        591 TGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGA 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGA 120

Qy        651 GAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGC 710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGC 180

Qy        711 GTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGC 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGC 240

Qy        771 AGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTT 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTT 300

Qy        831 CCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTGCC 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTGCC 360

Qy        891 TGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGC 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGC 420

Qy        951 TGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCC 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCC 480

Qy       1011 CTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGG 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTAGCTCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGG 540

Qy       1071 GCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCA 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCA 600

Qy       1131 CCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCT 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCT 660

Qy       1191 CCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCT 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGANAGGCAGAAGCCT 720

Qy       1251 GTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGG 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGG 780

Qy       1311 TGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGA 1370
              ||||||||||||||||||||||| |||  ||||||||||||||||||||||||||||||
Db        781 TGCTTATCAGGTGCCAAGCCCTGGTCTANGTGCTGGGAGTACAGCAGTGAGCAAAGGAGA 840

Qy       1371 CAATATTCCC--TGCTCACAGAGATGACAAACTGG 1403
              ||||||||||   |||  ||||||||||||||||
Db        841 CAATATTCCCCTGCCTCCAGGAGATGACAAACTGG 875


RESULT 10
BX325345
LOCUS       BX325345                945 bp    mRNA    linear   EST 02-MAY-2003
DEFINITION  BX325345 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
            Homo sapiens cDNA clone CS0DL004YF04 5-PRIME, mRNA sequence.
ACCESSION   BX325345
VERSION     BX325345.1  GI:30338449
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 945)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
```

```
           cgi-bin/cluster.cgi?seq=CS0AL004DC02QP1&cluster=10000.f. Contact :
           Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0AL004DC02QP1.
FEATURES             Location/Qualifiers
     source          1. .945
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DL004YF04"
                     /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                     25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      163 a    313 c    282 g    176 t     11 others
ORIGIN

  Query Match             40.8%;  Score 839.8;  DB 13;  Length 945;
  Best Local Similarity   97.8%;  Pred. No. 7.6e-68;
  Matches  860;  Conservative   10;  Mismatches    7;  Indels    2;  Gaps    2;

Qy         10 CCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACCCCTCTGGCT 69
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACCCCTCTGGCT 127

Qy         70 GCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTG-GATGACAACTTAGATAC 128
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        128 GCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGAGATGACAACTTAGATAC 187

Qy        129 CGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCT 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCCCCCTGCCT 247

Qy        189 GTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCG 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACTGCCTCCCG 307

Qy        249 TCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGGGCCCTGCA 308
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        308 TCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAGGCCCTGCA 367

Qy        309 CTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGT 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 CTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCCCTGGCCGT 427

Qy        369 GCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCT 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACTGAGGTCCT 487

Qy        429 GGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCT 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 GGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATGCACAGCCT 547

Qy        489 GGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGC 548
              |||||||| :|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 GGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGCCAGATGGC 606

Qy        549 CACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCG 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAGCTGTGTCG 666

Qy        609 CTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTG 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 CTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGACTCCTG 726

Qy        669 CGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGG 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 CGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTACGTGGG 786

Qy        729 ACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGGAGCCT 788
              ||||||||:|||||||||||||||||||:|||||||||||:| ||||||||||:||||||||
Db        787 ACCTGAGAWACTCAGCTCACGGGCCTYATACTCGGGCAARGGAGCCGATGTYTGGAGCCT 846

Qy        789 GGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAGCCTGT 848
              ||||||||||||||||||:||||||||||||||||||||||||||:|||||||||||||||
Db        847 GGGCGTGGCGCTCTTCACCAWGCTGGCCGGCCACTACCCCTTCCARGACTCGGAGCCTGT 906

Qy        849 CCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTT 887
              || |||||:||:|| ||:|||||||||| ||| |||||
Db        907 TCTTCTCTTYGGSAAAATYCGCCGCGGGGGCTAAGCCTT 945


RESULT 11
BX426485/c
LOCUS       BX426485               1201 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX426485 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
            clone CS0DH001YB11 3-PRIME, mRNA sequence.
ACCESSION   BX426485
VERSION     BX426485.1  GI:30770573
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0AH001CA06NP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0AH001CA06NP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DH001YB11"
                     /tissue_type="T CELLS (JURKAT CELL LINE)"
                     /cell_line="JURKAT CELL LINE"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
BASE COUNT      294 a    284 c    312 g    259 t     52 others
ORIGIN

  Query Match             40.5%;  Score 834;  DB 13;  Length 1201;
  Best Local Similarity   89.5%;  Pred. No. 2.2e-67;
  Matches  949;  Conservative   21;  Mismatches   81;  Indels    9;  Gaps    7;

Qy        926 GCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACC 985
              || |: |:  | :  :      |  | || | ||| ||    |||      |  || |  :
Db       1054 GCCGYTTYTGTTRSTKCTCCTCGTCGGAGCAGCTGACGCTAACACAACAGCATCTTCTGS 995

Qy        986 CCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTG 1045
              |    |||  |    :|    |||||||||| | |||||||||||||||||||:
Db        994 ACCCCTGGCGCGACAGAYCCGATCCCTTAGCCCAACCCCGATCCCATCTCTGGGAGS--T 937

Qy       1046 CCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAG 1105
              ||  | :|||||||:|  |:| ||||||||||||||||||||||||||||||||||||||
Db        936 GCCCAGAKGTCCCTGWTTTAYT-GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAG 878
```

```
Qy       1106 AAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGCTCAGCTGCCAACAGTGGATTGA 1165
             |||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        877 AAGT-GTTCTGTATGGCTA-GACCACCCTACTACACGCTCAGCTGCCAACAGTGGATTGA 820

Qy       1166 GTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTT 1225
             ||||||| |||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        819 GTTTGGGWGTAGCTCCAAGCCTTCTCCTGCCTCTGCACTGAGCCAAACCTTCAGTGCCTT 760

Qy       1226 CCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTC 1285
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 CCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTC 700

Qy       1286 CACACACATGCAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTG 1345
              |||||||  :|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TACACACAATSAGTTCCTGCTTGGGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTG 640

Qy       1346 GGAGTACAGCAGTGAGCAAAGGAGACAATATTCCCTGCTCACAGAGATGACAAACTGGCA 1405
             ||||:|||||||||||||:  |||||||||||: ||| ||||||||||||||||||||||
Db        639 GGAGWACAGCAGTGAGCAAMMCAGACAATATTCCM-GCTAACAGAGATGACAAACTGGCA 581

Qy       1406 TCCTTGAGCTGACAACACTTTTCCATGACCATAGGTCACTGTCTACACTGGGTACACTTT 1465
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        580 TCCTTGAGCTGACAACACTTTTCCATGACCATAGGTAACTGTCTACACTGGGTACACTTT 521

Qy       1466 GTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCC 1525
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        520 GTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCAAGGAAAATCCC 461

Qy       1526 TTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCC 1585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 TTTCACAAACAAACCAGCTGCCTTTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCC 401

Qy       1586 CTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTCTG 1645
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        400 CTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGACCCAAGCCCTGCTCACTCTG 341

Qy       1646 GGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAG 1705
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 GGTACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAG 281

Qy       1706 CCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAGTTCTTGTCTAACTCAA 1765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 CCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGCTAGTTCTTGTCTAACTCAA 221

Qy       1766 GACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAA 1825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAA 161

Qy       1826 GGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGG 1885
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||:|:
Db        160 GGTTGAGGGACAGGATTAGGCAGGGTTGTCCTGTGGCCACCTGGAAAGTCCCAGGTRGR 101

Qy       1886 ACTCTTCTGGGGACACTTGGGG--TCCACAATCCCAGGTCCATACTCTAGGTTTTGGATA 1943
             |||||||||||||||||||||    |||||||| ||||||||||||||||||||||||||
Db        100 ACTCTTCTGGGGACACTTGGGGATACCACAATCACAGGTCCATACTCTAGGTTTTGGATA 41

Qy       1944 C-CATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGA 1982
              | |||||||||||||||||  | |   |:|||::  ::|
Db         40 CNCATGAGTATGTATGTTTAATTTTWATTWATAWWTTWKA 1


RESULT 12
BX445733/c
LOCUS       BX445733               1201 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX445733 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
            clone CS0DH001YB11 3-PRIME, mRNA sequence.
ACCESSION   BX445733
VERSION     BX445733.1  GI:30782351
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS1DH001ZB03NP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS1DH001ZB03NP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DH001YB11"
                     /tissue_type=""T CELLS (JURKAT CELL LINE)"
                     /cell_line="JURKAT CELL LINE"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
BASE COUNT      263 a    292 c    303 g    245 t     98 others
ORIGIN

  Query Match           40.4%;  Score 832.8;  DB 13;  Length 1201;
  Best Local Similarity  87.7%;  Pred. No. 2.9e-67;
  Matches  933;  Conservative   32;  Mismatches   89;  Indels   10;  Gaps    7;

Qy        889 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 948
              | ||| ||:  ||  ::|||    ||  ||| ||:||:|||  | :| |  ||||||||:
Db       1077 CTTGCCGGMASCTTYSSCCCCYGCCHGTGTTTGKTTSGCT--CCYCCTTCTCGGGAGCCR 1020

Qy        949 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1008
              | | ||||: | ||||| ||| |:| |   |:| |||||| |::|: | : || || ::
Db       1019 G-TTAACGKTTAACAGCGACAAGMACCTCACWGAACCCCTTGYDGSCAWMAGAACCCRWT 961

Qy       1009 CCCTTAGCCCCAACCCGATCCCATCTCT---GGGAGGCTGCCCAGGTGGTCCCTGATGGA 1065
              |||||:|||| | ||    ||| | |   || ||:  :|||||||||||||:||:||
Db        960 GCCTTARCCCCCAACCCGATCCACATTTYTGGGAAGSTGRCCCAGGTGGTCCCTKATKGA 901

Qy       1066 CTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 1125
              : |||||:|||||||||||||||||||||||||:||||||:|||||||||||||:|||||
Db        900 AYTGGGCTKGACGAAGCCAGGGAAGAGGAGGGARACAGAGWAGTGGTTCTGTATKGCTAG 841

Qy       1126 GACCACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGC 1185
              ||| |||||||||:|||||| |||||| ||||||||||||||||||||||||||||||||
Db        840 GACAACCCTACTAMACGCTCTGCTGCCTACAGTGGATTGAGTTTGGGGGTAGCTCCAAGC 781

Qy       1186 CTTCTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGA 1245
              |||:||:|||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        780 CTTYTCYTGCCTCTGAACTGAGCCAAACCTTTAGTGCCTTCCAGAAGGGAGAAAGGCAGA 721

Qy       1246 AGCCTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGC 1305
              |||||||||||||||||||||||||||||||||| ||||||||| || ||||||||||||
Db        720 AGCCTGTGTGGAGTGTGCTGTGTACACTTCTGCTTTG-TCWACACACATGCAGTTCCTGC 662

Qy       1306 TTGGGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAA 1365
              ||||||||||:|||||||||||||||||||||||||||||||||| ||||||||||||||
Db        661 TTGGGTGCTTWTCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGCACAGCAGTGAGCAAA 602

Qy       1366 GGAGACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCC-TTGAGCTGACAACAC- 1423
              |||||||||||||| ||| |||||||||||||||||||||| ||| |||||| ||||
Db        601 GGAGACAATATTCCCNGCTNACAGAGATGACAAACTGGCATCCTTTGTGCTGACNACACT 542
```

```
Qy       1424 TTTTCCATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCC 1483
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTTCCATGACCATAGGTNACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCN 482

Qy       1484 ACTGATGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGC 1543
             ||||| ||||| ||| ||||||||||||||||||||||||||||||||||||||||||| ||
Db        481 ACTGANGCTGGNGCTNAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCGGC 422

Qy       1544 TGCCTTTGTATCTTGTACCTTTTCAGAGAAAGGGA-GGTATCCCTGTGCCAAAGGCTCCA 1602
              |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||:|||
Db        421 NGCCTTTGTATCTTGTACCTTTTCAGAGACAGGGACGGTATCCCTGTGCCAAAGGCYCCA 362

Qy       1603 GGCCTCTCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGC 1662
             |||||||||||||| |||||||||||||||| |||||||||||||||||||||||||||||
Db        361 GGCCTCTCCCCTGCNACTCAGGACCCAAGCCCCGCTCACTCTGGGAACTGTGTTCCCAGC 302

Qy       1663 ATCTCTGTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAG 1722
             ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCTCTGTCCTCTTGATTMAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAG 242

Qy       1723 AGATAAGAATCCAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAG 1782
             |||||||||||||| ||| ||||||| |||||||||| ||||||||||||||||||||||||
Db        241 AGATAAGAATCCANACTNTGAGGCTCGTTCTTGTCTCACTCAAGACTGTTCTGGAATGAG 182

Qy       1783 GGTCCAGGCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATT 1842
             |||||||||||||||||| |||||||||||||||||||||||||||||||| |||||||||
Db        181 GGTCCAGGCCTGTCAACCNTGGGGCTTCTGACCTGAGCACCAAGGTTGAGGNACAGGATN 122

Qy       1843 AGGCAGGGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACT 1902
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        121 AGGCAGGGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGNACTCTTCTGGGGACACT 62

Qy       1903 TGGGGTCCACAATCCCAGGTCCATACTCTAGGTTTTGGATACCA 1946
             |||||||||||||||||||||||||||||||          | | |
Db          61 TGGGGTCCACAATCCCAGGTCCATACTCTNNNNNNNNAAACCAA 18
```

RESULT 13
BG748142
LOCUS BG748142 863 bp. mRNA linear EST 15-MAY-2001
DEFINITION 602705453F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4842324 5', mRNA sequence.
ACCESSION BG748142
VERSION BG748142.1 GI:14058795
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 863)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1676 row: i column: 13
High quality sequence stop: 833.
FEATURES Location/Qualifiers
source 1. .863
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4842324"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. |"
BASE COUNT 135 a 300 c 264 g 164 t
ORIGIN

Query Match 39.8%; Score 818.6; DB 10; Length 863;
Best Local Similarity 98.2%; Pred. No. 6.6e-66;
Matches 838; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

```
Qy         38 GACGGGGCGAGATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGA 97
              | || ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 GGCGCGGCCAGATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGA 71

Qy         98 AGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAA 157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 AGCGGTTGGAGTTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAA 131

Qy        158 GTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAG 217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 GTGGGCCCCAGCCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAG 191

Qy        218 ATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCG 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATCGTGCAACTGCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCG 251

Qy        278 AGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCA 337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 AGGAGGGCGGGCGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCA 311

Qy        338 AGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGC 397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AGGTGTACCCCGTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGC 371

Qy        398 ACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTT 457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 ACAAGCATGTGGCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTT 431

Qy        458 TCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGC 517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 TCACTCGGACCCATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGC 491

Qy        518 CTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACG 577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 CTGAGGCTGCCGTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACG 551

Qy        578 GTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGA 637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GTCTGGTCCTGCGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGA 611

Qy        638 AGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT 697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 AGCTGGTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT 671

Qy        698 GGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCAT 757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 GGGACAAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCAT 731

Qy        758 ACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCG 817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 ACTCGGGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCG 791

Qy        818 GCCACTACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGG 877
              |||| |||||||||||||||||| |    |||||||||| |||||||| || ||||
Db        792 GCCAGTACCCCTTCCAGGACTCGGAAGCCTGTCTGCTCTTCGGGCAGATCCG-CGTGGGG 850

Qy        878 CCTACGCCTTGCC 890
```

```
              |||| ||||||||
Db        851 CCTAAGCCTTGCC 863


RESULT 14
AL556690
LOCUS       AL556690                1201 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL556690 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK005YC05 5-PRIME, mRNA sequence.
ACCESSION   AL556690
VERSION     AL556690.2  GI:31278491
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12899595.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DK005AB03QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK005AB03QP1.
FEATURES              Location/Qualifiers
     source           1. .1201
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DK005YC05"
                      /cell_type="HELA CELLS COT 25-NORMALIZED"
                      /cell_line="HELA"
                      /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      205 a    366 c    336 g    232 t     62 others
ORIGIN

  Query Match           39.7%;  Score 818.4;  DB 9;  Length 1201;
  Best Local Similarity   96.2%;  Pred. No. 5.8e-66;
  Matches  841;  Conservative   15;  Mismatches   16;  Indels    2;  Gaps    2;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 197

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 257

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 317

Qy        181 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CCCTGCCTGTTGCCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 377

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 437

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 497

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 557

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 617

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              |||||||||||||||| :|||||||||||||||||||||||||||||||||||||||||
Db        618 CACAGCCTGGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 676

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 736

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        737 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTAGAGAACCTGGAG 796

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 856

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 916

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||:|||||||||||||||:|:||::|||||||| ||||||||||||:| |||||:
Db        917 TGGAGCCTKGGCGTGGCGCTCTTYAMCAWKCTGGCCGGGCACTACCCCTTCMA-GACTCS 975

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCG 874
              |::|:|:|| :  : |  |  ||    ||: | |
Db        976 GRRCYTKTCTKCTYTTCGGMAAATCSGCGSGGGG 1009


RESULT 15
BX443303
LOCUS       BX443303                948 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  BX443303 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG007YF05 5-PRIME, mRNA sequence.
ACCESSION   BX443303
VERSION     BX443303.1  GI:31018564
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10000.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DG007CC03QP1&cluster=10000.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DG007CC03QP1.
FEATURES              Location/Qualifiers
     source           1. .948
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
```

```
                     /db_xref="taxon:9606"
                     /clone="CS0DG007YP05"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
BASE COUNT      152 a    326 c    296 g    169 t      5 others
ORIGIN

  Query Match             39.7%;  Score 816.6;  DB 13;  Length 948;
  Best Local Similarity   99.4%;  Pred. No. 9.6e-66;
  Matches  838;  Conservative    2;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 167

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 227

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 287

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 347

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        348 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 407

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        408 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCCGTCCAGGAAGCC 466

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 526

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 586

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              |||||||||||||||| :|||||||||||||||||||||||||||||||||||||||||
Db        587 CACAGCCTGGTGCGAA-SCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 645

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 705

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 765

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 825

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 885

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        886 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCARGACTCG 945
```

```
Qy        841 GAG 843
              |||
Db        946 GAG 948
```

Search completed: January 16, 2004, 16:46:14
Job time : 4628.16 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:09:04 ; Search time 583.591 Seconds
(without alignments)
9524.031 Million cell updates/sec

Title: US-09-909-474D-1
Perfect score: 2059
Sequence: 1 gctctgagccccggcggcgc.....aaaaaaaaaaaaaaaaaaaa 2059

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1999.2	97.1	2116	22	AAF30480	Human protein phos
2	1997.2	97.0	2059	22	AAI59850	Human polynucleoti
3	1997.2	97.0	2092	22	AAI58064	Human polynucleoti
4	1767.8	85.9	2048	21	AAC77866	Human cancer assoc
5	1073.8	52.2	1077	22	AAS06709	Polynucleotide seq
6	1054.6	51.2	1083	24	ABN86479	Human tribbles hom
7	1004.4	48.8	1085	21	AAZ61155	cDNA JJ503-KS enco
8	802.2	39.0	972	22	AAH76218	Human kinase PKIN-
9	645.2	31.3	1076	24	ABL39747	Human NS cDNA sequ
10	645.2	31.3	1076	24	ABL39762	Human NS cDNA sequ
c 11	562.4	27.3	996	23	AAS91231	DNA encoding novel
12	533	25.9	541	24	ABK55110	Human colon cancer
c 13	505.2	24.5	2562	16	AAQ89817	DNA encoding the A
c 14	426	20.7	426	24	ABL87304	Human ovarian canc
15	330.2	16.0	396	22	AAS37461	Novel human diagno
c 16	315.4	15.3	353	24	ABL62004	Colon adenocarcino
17	296	14.4	534	23	AAS68600	DNA encoding novel
18	274.2	13.3	3319	24	ABN86478	Human tribbles hom
19	268.8	13.1	562	21	AAC78131	Human cancer assoc
20	263.2	12.8	1909	22	AAS45035	cDNA encoding nove
21	251.4	12.2	3280	25	ABX63269	Human cDNA #269 di
22	247	12.0	3324	24	AAS94913	Human DNA sequence
23	242.6	11.8	690	24	ABT09024	Phase-1 Rat CT gen
24	234	11.4	675	24	ABK83550	Human cDNA differe
25	233.4	11.3	942	25	ABX74428	Human cDNA sequenc
26	232.4	11.3	818	25	ACC45125	Human C8FW encodin
27	222.2	10.8	621	25	ACC45126	Human C8FW encodin
28	207.6	10.1	1943	22	AAS45223	cDNA encoding nove
29	205	10.0	207	21	AAC03717	Human secreted pro
c 30	165	8.0	498	24	ABK63084	Rat sequence diffe
31	163.6	7.9	7282	22	AAK86318	Human immune/haema
c 32	158	7.7	562	22	ABA61922	Human foetal liver
c 33	158	7.7	562	22	AAK10231	Human brain expres
c 34	158	7.7	562	22	AAK36130	Human bone marrow
c 35	158	7.7	562	22	AAI41844	Probe #10530 used
c 36	158	7.7	562	23	ABS35823	Human liver single
c 37	158	7.7	562	24	ABS10267	Human genome-deriv
38	158	7.7	25772	22	AAK81332	Human immune/haema
c 39	156	7.6	417	22	ABA74417	Human foetal liver
c 40	156	7.6	417	22	AAK22879	Human brain expres
c 41	156	7.6	417	22	AAK49052	Human bone marrow
c 42	156	7.6	417	22	AAI54879	Probe #23565 used
c 43	156	7.6	417	23	ABS48709	Human liver single
c 44	156	7.6	417	24	ABS22668	Human genome-deriv
45	142.8	6.9	2038	23	ABL29125	Drosophila melanog

ALIGNMENTS

RESULT 1
AAF30480
ID AAF30480 standard; cDNA; 2116 BP.
XX
AC AAF30480;
XX
DT 29-MAY-2001 (first entry)
XX
DE Human protein phosphatase and kinase protein-5 cDNA 1271505CB1.
XX
KW Protein phosphatase and kinase protein; PPHKP-5; human;
KW gastrointestinal disorder; immune system disorder;
KW neurological disorder; cell proliferative disorder; cancer;
KW diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 161..1237
FT /*tag= a
FT misc_feature 326..385
FT /*tag= b
FT /note= "unique fragment"
XX
PN WO200120004-A2.
XX
PD 22-MAR-2001.
XX
PF 14-SEP-2000; 2000WO-US25515.
XX
PR 15-SEP-1999; 99US-0154141.

XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y;
PI Lu DAM;
XX
DR WPI; 2001-244811/25.
DR P-PSDB; AAB20326.
XX
PT Novel human protein phosphatase and kinase proteins for diagnosis,
PT treatment and prevention of gastrointestinal, immune system,
PT neurological and cell proliferative disorders -
XX
PS Claim 5; Page 98; 103pp; English.
XX
CC The present sequence is that of cDNA encoding novel human
CC protein phosphatase and kinase protein PPHKP-5 (see AAB20326).
CC The cDNA was initially identified in Incyte Clone ID No. 1271505CB1,
CC from a foetal lung tissue library. The PPHKP-5 gene was localised
CC to human chromosome 20. Tissues that express PPHKP-5 (as a
CC fraction of total tissues expressing PPHKP-5) include reproductive
CC (0.288), gastrointestinal (0.212) and haematopoietic or immune
CC (0.192). Diseases or conditions associated with tissues expressing
CC PPHKP-5 (as a fraction of total tissues expressing PPHKP-5) include
CC cancer (0.577), inflammation or trauma (0.327) and cell
CC proliferation (0.308). The encoded protein shows homology to rat
CC kinase. The invention provides human PPHKP-1 to -11 polypeptides
CC (see AAB20322-32) and polynucleotides (see AAP30476-86). It also
CC provides expression vectors, host cells, antibodies, agonists
CC and antagonists, as well as methods for diagnosing, treating or
CC preventing disorders associated with expression of PPHKP, including
CC gastrointestinal disorders, immune system disorders, neurological
CC disorders and cell proliferative disorders, including cancer.
XX
SQ Sequence 2116 BP; 407 A; 663 C; 609 G; 437 T; 0 other;

Query Match 97.1%; Score 1999.2; DB 22; Length 2116;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 172

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 232

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 292

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 352

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        353 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 412

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 472

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 532

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 592

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        593 CACAGCCTGGTGCGAACGCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 652

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 712

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 772

Qy        661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 832

Qy        721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 892

Qy        781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 952

Qy        841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 1012

Qy        901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1072

Qy        961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1132

Qy       1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1192

Qy       1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1252

Qy       1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1312

Qy       1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1372

Qy       1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1432

Qy       1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1492

Qy       1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1552

Qy       1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1612

Qy       1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1672

Qy       1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1673 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1732

Qy       1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1792

Qy       1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1852

Qy       1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1912

Qy       1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1972

Qy       1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 2032

Qy       1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2092

Qy       1981 GAATTATGAAATAAAAAAAAAAAA 2004
              ||||||||||||||||||||||||
Db       2093 GAATTATGAAATAAAAAAAAAAAA 2116
```

RESULT 2
AAI59850
ID AAI59850 standard; cDNA; 2059 BP.
XX
AC AAI59850;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 3839.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM40694.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 3839; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 2059 BP; 400 A; 645 C; 583 G; 431 T; 0 other;

Query Match 97.0%; Score 1997.2; DB 22; Length 2059;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 117

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 177

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 237

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 297

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        298 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 357

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        358 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAGGGTGTACCCCGTCCAGGAAGCC 417

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        418 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 477

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 537

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 597

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 657

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 717
```

```
Qy       661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       718 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 777

Qy       721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       778 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 837

Qy       781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       838 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 897

Qy       841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       898 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 957

Qy       901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       958 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1017

Qy       961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1018 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1077

Qy      1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1078 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1137

Qy      1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1138 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1197

Qy      1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1198 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1257

Qy      1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1258 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1317

Qy      1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1318 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1377

Qy      1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1378 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1437

Qy      1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1438 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1497

Qy      1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1498 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1557

Qy      1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1558 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1617

Qy      1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1618 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1677

Qy      1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1678 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1737

Qy      1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1738 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1797

Qy      1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1798 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1857

Qy      1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1858 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1917

Qy      1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1918 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1977

Qy      1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1978 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2037

Qy      1981 GAATTATGAAATAAAAAAAAAA 2002
            ||||||||||||||||||||||
Db      2038 GAATTATGAAATAAAAAAAAAA 2059
```

RESULT 3
AAI58064
ID AAI58064 standard; cDNA; 2092 BP.
XX
AC AAI58064;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 267.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM38908.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 267; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 2092 BP; 405 A; 658 C; 593 G; 436 T; 0 other;

Query Match 97.0%; Score 1997.2; DB 22; Length 2092;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy    1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 150

Qy   61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  151 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 210

Qy  121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  211 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 270

Qy  181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  271 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 330

Qy  241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db  331 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 390

Qy  301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  391 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAGGGTGTACCCCGTCCAGGAAGCC 450

Qy  361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  451 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 510

Qy  421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  511 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 570

Qy  481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  571 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 630

Qy  541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  631 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 690

Qy  601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCT[illegible]GCTGGAGAACCTGGAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  691 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 750

Qy  661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  751 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 810

Qy  721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  811 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 870

Qy  781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  871 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 930

Qy  841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  931 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 990

Qy  901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  991 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1050

Qy  961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1051 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCA 1110

Qy 1021 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1111 ACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA 1170

Qy 1081 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1171 GCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACA 1230

Qy 1141 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1231 CGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTG 1290

Qy 1201 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1291 AACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGTGTGGAGTG 1350

Qy 1261 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1351 TGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTGGGTGCTTATCAG 1410

Qy 1321 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1411 GTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGAGACAATATTCCC 1470

Qy 1381 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1471 TGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTCCATGACCATAGG 1530

Qy 1441 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1531 TCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGATGCTGGTGCTCA 1590

Qy 1501 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1591 GGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCTTTGTATCTTGTA 1650

Qy 1561 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1651 CCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACT 1710

Qy 1621 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1711 CAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATT 1770

Qy 1681 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1771 AAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTA 1830

Qy 1741 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1831 TGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACC 1890

Qy 1801 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1891 ATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGT 1950

Qy 1861 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1951 GGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAG 2010
```

```
Qy       1921 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 GTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGA 2070

Qy       1981 GAATTATGAAATAAAAAAAAAA 2002
             |||||||||||||||||||||||
Db       2071 GAATTATGAAATAAAAAAAAAA 2092


RESULT 4
AAC77866
ID   AAC77866 standard; cDNA; 2048 BP.
XX
AC   AAC77866;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:260.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43657.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 825-826; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 2048 BP; 404 A; 619 C; 585 G; 434 T; 6 other;

  Query Match            85.9%;  Score 1767.8;  DB 21;  Length 2048;
  Best Local Similarity  95.9%;  Pred. No. 1.1e-303;
  Matches 1882;  Conservative    4;  Mismatches    1;  Indels   76;  Gaps    4;

Qy         49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 215

Qy        109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 275

Qy        169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
              ||||||||||||||||||||||||||||||
Db        276 CCCAGACTGCCCCCCTGCCTGTTGCCCCT------------------------------ 304

Qy        229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
                                                  ||||||||||||||||||||
Db        305 ------------------------------------------GAGCCCGAGGAGGGCGGG 322

Qy        289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
              |||||||||| || |||||||||||||||||||||||||||||||||||||| |||||||||
Db        323 CGGGCCTACCAGG-CCTGCACTGCCCTACAGGCACTGAGTATACCTGCAA-GTGTACCCC 380

Qy        349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        381 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTAYGCGCGGCTGCCCCCGCACAAGCATGTG 440

Qy        409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 500

Qy        469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        501 CATGGGGACATGCACAGCCTGGTGCGAAG-CGCCACCGTATCCCTGAGCCTGAGGCTGCC 559

Qy        529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 619

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 679

Qy        649 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 739

Qy        709 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 768
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        740 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCAMGGGCCTCATACTCGGGCAAG 799

Qy        769 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 859

Qy        829 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 919

Qy        889 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 979

Qy        949 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1039

Qy       1009 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1068
              ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1040 CCCTTAGCYCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1099

Qy       1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGAC 1128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGAC 1159

Qy       1129 CACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTT 1188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1160 CACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTT 1219

Qy       1189 CTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGC 1248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1220 CTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGC 1279

Qy       1249 CTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTG 1308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1280 CTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTG 1339

Qy       1309 GGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGA 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1340 GGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGA 1399

Qy       1369 GACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTC 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1400 GACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTC 1459

Qy       1429 CATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGA 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1460 CATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGA 1519

Qy       1489 TGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCT 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1520 TGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCT 1579

Qy       1549 TTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1580 TTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 1639

Qy       1609 TCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCT 1668
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       1640 TCCCCTGCAACTCAGGACCCAAGCCCAGCTCACTCTGGGAACTGTRTTCCCAGCATCTCT 1699

Qy       1669 GTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAA 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1700 GTCCTCTTGATTAAGAGATTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAA 1759

Qy       1729 GAATCCAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCA 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1760 GAATCCAAACTATGAGGCTAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCA 1819

Qy       1789 GGCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAG 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1820 GGCCTGTCAACCATGGGGCTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAG 1879

Qy       1849 GGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGT 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1880 GGTCTGTCCTGTGGCCACCTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGT 1939

Qy       1909 CCACAATCCCAGGTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTG 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1940 CCACAATCCCAGGTCCATACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTG 1999

Qy       1969 CCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAAA 2011
             |||||||||||||||||||||||||||||||||||||||||||
Db       2000 CCTAATAAAGGAGAATTATGAAATAAAAAAAAAAAAAAAAAA 2042


RESULT 5
AAS06709
ID   AAS06709 standard; cDNA; 1077 BP.
XX
AC   AAS06709;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Polynucleotide sequence encoding human protein kinase #9.
XX
KW   Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW   metabolic disorder; immune related disease; neurological disorder;
KW   neurodegenerative disorder; inflammatory disorder; infectious disease;
KW   reproductive disorder; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200138503-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US32085.
XX
PR   24-NOV-1999;   99US-0167482.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Whyte D,  Manning G,  Sudarsanam S,  Martinez R;
PI   Flanagan P,  Clary D;
XX
DR   WPI; 2001-343950/36.
DR   P-PSDB; AAU03509.
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections -
XX
PS   Example 1; Figure 1; 433pp; English.
XX
CC   AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC   novel protein kinases have been identified as members of the tyrosine
CC   or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be
CC   used for gene therapy and as DNA probes in diagnostic assays.
CC   The protein kinase polypeptides may be used as antigens in the production
CC   of antibodies against the protein kinases and in assays to identify
CC   modulators of protein kinase expression and activity.
XX
SQ   Sequence 1077 BP; 174 A; 373 C; 325 G; 205 T; 0 other;

  Query Match           52.2%;  Score 1073.8;  DB 22;  Length 1077;
  Best Local Similarity   99.8%;  Pred. No. 6e-181;
  Matches 1075;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy        109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240
```

```
Qy        289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTG 360

Qy        409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 600

Qy        649 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 660

Qy        709 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 720

Qy        769 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 780

Qy        829 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTG 840

Qy        889 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 900

Qy        949 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 960

Qy       1009 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1020

Qy       1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 1077
```

```
RESULT 6
ABN86479
ID   ABN86479 standard; DNA; 1083 BP.
XX
AC   ABN86479;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human tribbles homologue-3 (htrb-3) polypeptide encoding DNA.
XX
KW   Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic; osteopathic;
KW   opthalamological; cardiant; cytostatic; haemostatic; immunosuppressive;
KW   antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
KW   tumour necrosis factor; TNF; gene; htrb-3; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1083
FT                   /*tag= a
FT                   /product= "htrb-3"
FT                   /note= "tribbles polypeptide"
XX
PN   WO200253743-A2.
XX
PD   11-JUL-2002.
XX
PF   08-JAN-2002; 2002WO-US00070.
XX
PR   08-JAN-2001; 2001US-260294P.
XX
PA   (INTE-) INTERLEUKIN GENETICS INC.
XX
PI   Dower S,  Quanstrom E,  Kiss-Toth E;
XX
DR   WPI; 2002-590635/63.
DR   P-PSDB; ABB80976.
XX
PT   Novel isolated human tribbles homologue-1 polypeptide for inhibiting
PT   AP-1-mediated inflammatory signal in a cell, and activating
PT   ERK-mediated signal e.g. AP-1-mediated gene activation signal in a cell
PT   -
XX
PS   Example 10; Fig 11A; 131pp; English.
XX
CC   The invention provides an isolated human tribbles homologue-1 (htrb-1,
CC   also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC   htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC   signal in a cell. The polypeptide employed in the method is preferably
CC   htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or
CC   htrb-3 N C. It is also useful for providing htrb agonist activity for
CC   activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC   signal, an estrogen receptor-mediated gene activation signal, an
CC   fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
CC   in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC   inflammatory signal in a cell such as tumor necrosis factor (TNF)
CC   induced inflammatory signal, or an interleukin induced inflammatory
CC   signal. htrb proteins are useful in screening assays, predictive medicine
CC   and in therapeutics or prophylactics. The htrb proteins are useful for
CC   screening compounds e.g. for treating and/or preventing diseases caused
CC   by abnormal htrb activity, such as rheumatoid arthritis, diabetes,
CC   psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction
CC   and cancers. The htrb therapeutics are useful for antagonizing
CC   interleukin-1 dependent disorders of human placenta, intraventricular
CC   hemorrhage, neonatal white matter damage and subsequent cerebral palsy;
CC   and inflammation or autoimmune disorders. The present sequence represents
CC   a DNA encoding the htrb-3 polypeptide.
XX
SQ   Sequence 1083 BP; 176 A; 374 C; 327 G; 206 T; 0 other;

  Query Match             51.2%;  Score 1054.6;  DB 24;  Length 1083;
  Best Local Similarity   99.1%;  Pred. No. 1.5e-177;
  Matches 1073;  Conservative    0;  Mismatches    4;  Indels    6;  Gaps    1;

Qy         49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy        109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAG 120

Qy        169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy        229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy        289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy        349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
              |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||
Db        301 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTATGCGCGGGTGCCCCCGCACAAGCATGTG 360

Qy        409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 420

Qy        469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 480

Qy        529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 540

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGC------TGACCGTGAGAGGAAGAAGCTG 642
              |||||||||||||||||||||||||||||||      |||||||||| ||||||||||
Db        541 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGACCGTGAGAAGAAGAAGCTG 600

Qy        643 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTGCTGGAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGAC 660

Qy        703 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGCACGCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCG 720

Qy        763 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGCAAGGCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCAC 780

Qy        823 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTAC 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACCCCTTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTAC 840

Qy        883 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCCTTGCCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGG 900

Qy        943 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGCCAGCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGAC 960

Qy       1003 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 CCGATGCCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGAT 1020

Qy       1063 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGACTGGGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGC 1080

Qy       1123 TAG 1125
              |||
Db       1081 TAG 1083
```

```
RESULT 7
AAZ61155
ID   AAZ61155 standard; cDNA; 1085 BP.
XX
AC   AAZ61155;
XX
DT   30-MAY-2000  (first entry)
XX
DE   cDNA JJ503-KS encoding domains VIA to XI of a protein kinase.
XX
KW   Kinase activity; molecular weight marker; isoelectric focusing marker;
KW   peptide fragmentation control; cellular signal transduction; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..1081
FT                   /*tag= a
FT                   /note= "partial sequence"
XX
PN   WO200008180-A2.
XX
PD   17-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US17630.
XX
PR   04-AUG-1998;   98US-0095270.
PR   11-SEP-1998;   98US-0099972.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Virca GD,  Bird TA,  Anderson DM,  Marken JS;
XX
DR   WPI; 2000-195584/17.
DR   P-PSDB; AAY69157.
XX
PT   New human kinase polypeptides and polynucleotides used as molecular
PT   weight markers and as controls for peptide fragmentation  -
XX
PS   Claim 1; Page 7; 60pp; English.
XX
CC   The present sequence encodes a partial polypeptide which has kinase
CC   activity. The kinase polynucleotides can be used to express the
CC   polypeptides, and as probes to identify nucleic acids encoding
CC   proteins having kinase activity. The kinase polypeptides and
CC   fragmented polypeptides are used as molecular weight and isoelectric
CC   focusing markers, and as controls for peptide fragmentation. They also
CC   have a number of therapeutic uses as kinases play a central role in
CC   cellular signal transduction. The polypeptides could also be used to
CC   identify binding partner proteins. The polypeptides can also be used as
CC   a reagent to identify any proteins that the polypeptide regulates, and
CC   proteins with which it might interact. The polypeptides may also be
CC   used for preparation of antibodies. The antibodies can be used in
CC   assays to detect the presence of the protein, and to purify the protein
CC   by immunoaffinity chromatography.
XX
SQ   Sequence 1085 BP; 164 A; 389 C; 338 G; 194 T; 0 other;

  Query Match             48.8%;  Score 1004.4;  DB 21;  Length 1085;
  Best Local Similarity   99.9%;  Pred. No. 1.1e-168;
  Matches 1005;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 139

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 199

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 259

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 319

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        320 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 379

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          380 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 439

Qy          361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          440 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 499

Qy          421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          500 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 559

Qy          481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          560 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 619

Qy          541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          620 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 679

Qy          601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          680 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 739

Qy          661 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          740 GACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCC 799

Qy          721 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          800 TACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTC 859

Qy          781 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          860 TGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCG 919

Qy          841 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          920 GAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGCCTACGCCTTGCCTGCAGGCCTC 979

Qy          901 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          980 TCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTC 1039

Qy          961 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1006
                 ||||||||||||||||||||||||||||||||||||||||||||||
Db         1040 ACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGA 1085
```

RESULT 8
AAH76218

```
ID   AAH76218 standard; cDNA; 972 BP.
XX
AC   AAH76218;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Human kinase PKIN-10 encoding cDNA.
XX
KW   PKIN; kinase; cytostatic; immunosuppressive; immunostimulant; human;
KW   antiarteriosclerotic; cardiant; gene therapy; antisense therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..972
FT                   /*tag=  a
FT                   /product= "PKIN-10"
XX
PN   WO200160991-A2.
XX
PD   23-AUG-2001.
XX
PF   16-FEB-2001; 2001WO-US05240.
XX
PR   17-FEB-2000; 2000US-0183682.
PR   02-MAR-2000; 2000US-0186559.
PR   09-MAR-2000; 2000US-0188606.
PR   17-MAR-2000; 2000US-0189998.
PR   30-MAR-2000; 2000US-0193851.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Buford N,  Gandhi AR,  Patterson C,  Khan FA,  Yue H;
PI   Hafalia A,  Shih LL,  Tribouley CM,  Yao MG,  Burrill JD,  Marcus GA;
PI   Zingler KA,  Lu DAM,  Bandman O,  Policky JL,  Griffin JA,  Thornton M;
PI   Nguyen DB,  Lal P,  Walsh RT;
XX
DR   WPI; 2001-514771/56.
DR   P-PSDB; AAB85791.
XX
PT   isolated human kinase polypeptides useful in the diagnosis, treatment
PT   and prevention of cancer, immune disorders and disorders affecting
PT   growth and development -
XX
PS   Claim 5; Page 125; 126pp; English.
XX
CC   The invention provides human kinases (PKIN) and polynucleotides encoding
CC   PKIN. The PKIN polypeptides can be expressed using standard recombinant
CC   methodology. The PKIN polypeptides, polynucleotides, modulators and
CC   specific antibodies are useful in the diagnosis, treatment and prevention
CC   of cancer, immune disorders, disorders affecting growth and development,
CC   atherosclerosis, and other cardiovascular diseases, and lipid disorders
CC   and in the assessment of the effects of exogenous compounds on the
CC   expression of nucleic acid sequences of human kinases. The present
CC   sequence represents a cDNA encoding a human PKIN-10 polypeptide.
XX
SQ   Sequence 972 BP; 167 A; 329 C; 294 G; 182 T; 0 other;

  Query Match             39.0%;  Score 802.2;  DB 22;  Length 972;
  Best Local Similarity   87.7%;  Pred. No. 6.7e-133;
  Matches  944;  Conservative    0;  Mismatches   28;  Indels  105;  Gaps    2;

Qy           49 ATGCGAGCCACCCCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 108
                 ||||||||||||||| |||||||| |||||||||||||||||||||||||||||||||||
Db            1 ATGCGAGCCACCCCTTTGGCTGCTTCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAG 60

Qy          109 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGCCCCCAG 168
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 TTGGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGCCCCCAG 120

Qy          169 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 228
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 CCCAGACTGCCCCCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACT 180

Qy          229 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 288
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GCTGTGGCCACTGCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGG 240

Qy          289 CGGGCCTACCGGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 348
                 |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 CGGGCCTACCAGGCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCC 300

Qy          349 GTCCAGGAAGCCCTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTG 408
                 |||||||||||||||||||||||||| ||||| ||||||||||||||| |||||||||||
Db          301 GTCCAGGAAGCCCTGGCCGTGCTGGAACCCTATGCGCGGCTGCCCCCGGACAAGCATGTG 360

Qy          409 GCTCGGCCCACTGAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACC 468
                 ||| |||||||||||| ||||||||||||||| || |||||||||||||||||| |||||
Db          361 GCTTGGCCCACTGAGGGCCTGGCTGGTACCCAACTTCTCTACGCCTTTTTCACTTGGACC 420

Qy          469 CATGGGGACATGCACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCC 528
                 ||||||||||||||||| ||| |  |
Db          421 CATGGGGACATGCACAGGCTGATAGG---------------------------------- 446

Qy          529 GTGCTCTTCCGCCAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTG 588
                                  |  |||  ||    ||||||||| |||||
```

```
Db        447 ----------------GCACACGCCCTGTGCGCACTGTGACCAG---------------- 474

Qy        589 CGTGATCTCAAGCTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTG 648
                                                  ||||||||||||||||||
Db        475 ------------------------------------ACACGGAAGAAGCTGGTGCTG 495

Qy        649 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 GAGAACCTGGAGGACTCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCAC 555

Qy        709 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 GCGTGCCCAGCCTACGTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAG 615

Qy        769 GCAGCCGATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 GCAGCCAATGTCTGGAGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCC 675

Qy        829 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 TTCCAGGACTCGGAGCCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTG 735

Qy        889 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 CCTGCAGGCCTCTCGGCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCA 795

Qy        949 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GCTGAACGGCTCACAGCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATG 855

Qy       1009 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 CCCTTAGCCCCAACCCGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTG 915

Qy       1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAG 972
```

RESULT 9
ABL39747
ID ABL39747 standard; cDNA; 1076 BP.
XX
AC ABL39747;
XX
DT 10-MAY-2002 (first entry)
XX
DE Human NS cDNA sequence SEQ ID NO:57.
XX
KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200206315-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-IL00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Freilich S, Bernstein J;
XX
DR WPI; 2002-155037/20.
DR P-PSDB; ABB06093.
XX
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 1; Page 107; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SQ Sequence 1076 BP; 172 A; 352 C; 346 G; 194 T; 12 other;

  Query Match   31.3%; Score 645.2; DB 24; Length 1076;
  Best Local Similarity 98.2%; Pred. No. 4e-105;
  Matches 670; Conservative   3; Mismatches   6; Indels   3; Gaps   2;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 340

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 400

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 460

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 520

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        521 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 580

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 640

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        641 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 700

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 760
```

```
Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              ||||||||||||||||||::||||||||||||||||||||||||||||||||||||||||
Db        761 CACAGCCTGGTGCGAAGCSSCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 820

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 880

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              |||||||||||||||||||||||||||||||| |||||||||||||||||  |  |
Db        881 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG-AGAAGCTGGTGCTGGAGA--CCTGGA 937

Qy        661 GACTCCTGCGTGCTGACTGGGC 682
              |||||:||||||||||||||||
Db        938 GACTCSTGCGTGCTGACTGGGC 959
```

RESULT 10
ABL39762
ID ABL39762 standard; cDNA; 1076 BP.
XX
AC ABL39762;
XX
DT 10-MAY-2002 (first entry)
XX
DE Human NS cDNA sequence SEQ ID NO:72.
XX
KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200206315-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-IL00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Freilich S, Bernstein J;
XX
DR WPI; 2002-155037/20.
DR P-PSDB; ABB06108.
XX
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 1; Page 117; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SQ Sequence 1076 BP; 172 A; 352 C; 346 G; 194 T; 12 other;

Query Match 31.3%; Score 645.2; DB 24; Length 1076;
Best Local Similarity 98.2%; Pred. No. 4e-105;
Matches 670; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 340

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 400

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 460

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 520

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        521 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCAG 580

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGTACCCCGTCCAGGAAGCC 640

Qy        361 CTGGCCGTGCTGGAGCCCTACGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 420
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        641 CTGGCCGTGCTGGAGCCCTATGCGCGGCTGCCCCCGCACAAGCATGTGGCTCGGCCCACT 700

Qy        421 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 GAGGTCCTGGCTGGTACCCAGCTCCTCTACGCCTTTTTCACTCGGACCCATGGGGACATG 760

Qy        481 CACAGCCTGGTGCGAAGCCGCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 540
              |||||||||||||||||::||||||||||||||||||||||||||||||||||||||||||
Db        761 CACAGCCTGGTGCGAAGCSSCCACCGTATCCCTGAGCCTGAGGCTGCCGTGCTCTTCCGC 820

Qy        541 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 CAGATGGCCACCGCCCTGGCGCACTGTCACCAGCACGGTCTGGTCCTGCGTGATCTCAAG 880

Qy        601 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAG 660
              |||||||||||||||||||||||||||||||| |||||||||||||||||  |  |
Db        881 CTGTGTCGCTTTGTCTTCGCTGACCGTGAGAGG-AGAAGCTGGTGCTGGAGA--CCTGGA 937

Qy        661 GACTCCTGCGTGCTGACTGGGC 682
              |||||:||||||||||||||||
Db        938 GACTCSTGCGTGCTGACTGGGC 959
```

RESULT 11
AAS91231/c
ID AAS91231 standard; cDNA; 996 BP.

```
XX
AC   AAS91231;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #27035.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG27044.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 27035; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 996 BP; 248 A; 244 C; 251 G; 249 T; 4 other;

  Query Match           27.3%;  Score 562.4;  DB 23;  Length 996;
  Best Local Similarity  88.1%;  Pred. No. 1.8e-90;
  Matches  860;  Conservative    0;  Mismatches   71;  Indels   45;  Gaps   21;

Qy       1073 TGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACC 1132
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 TCGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACC 917

Qy       1133 CTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCC 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 CTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCC 857

Qy       1193 TGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGT 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        856 TGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGCCTGT 797

Qy       1253 GTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACAT--GCAGTTCCTGCTTGGG 1310
              ||||||||||||||||||||||||||||||||||||||||||    ||||||||||||||
Db        796 GTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATTGCAAGTTCCTGCTTGGG 737

Qy       1311 TGCTTATCAGGTGCC-AAGCCCTGTTCTCGGTGCTGGG-AGTACAGC-AGTGAGCAAAGG 1367
              ||||||||||||||| |||||||||||||||||||||| |||||||| ||||||||||||
Db        736 TGCTTATCAGGTGCCAAAGCCCTGTTCTCGGTGCTGGGAAGTACAGCAAGTGAGCAAAGG 677

Qy       1368 AGACAATATTCCCTGCTCACAGAGAT----GACAAACTGGCATCCTTGAGCTGACAACAC 1423
              |||||||||||||||||||||||||     | |   |   | ||||||||||||||||
Db        676 AGACAATATTCCCTGCTCACAGAGATTGACAAAACTGGGCATCCTTTGAGCTGACAACAC 617

Qy       1424 TTTTCCAT-GACCATAGGTCAC-TGTCTACACTGGG--TACACTTTGTACCAGTGTCGGC 1479
              |||||||| ||||||||||||| |||||||||||||       ||||||||||||||||
Db        616 TTTTCCATGGACCATAGGTCACTTGTCTACACTGGGGTACACTTTTGTACCAGTGTCGGC 557

Qy       1480 CTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAAC 1539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 CTCCACTGATGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAAC 497

Qy       1540 CAG-CTGCCTTTGTATCTTG---TACCTTTTCAGAGAAA-GGGAGGTATCCCTGTGCCAA 1594
              ||| |||||||||||||||     | |||||||||||| |||||||||||||| | ||||
Db        496 CAGCCTGCCTTTGTATCTTGGTACCCTTTTTCAGAGAAAGGGGAGGTATCCCTTTTCCAA 437

Qy       1595 AGGCTCCAGGCCTCTCCCCTGCAACTCA-------GGACCCAAGCCCAGCTC-ACTCTGG 1646
              |||||||| | |||| | | | |        |  ||||| || ||||| |||||||
Db        436 AGGCTCCAAGGCCCTCCTCCCCTTCCAAATTCAGGGACCCCAACCCAAGCTCAACTCTGG 377

Qy       1647 GAACTGTGTTCCCAGCA---TCTCTGTCCTCTTGATTAA--GAGATTCTCCTTCCAGGCC 1701
              || || ||||||||||    |||     |  ||||||||   ||||||||||||||| |
Db        376 GACCTTTGTTCCCAGCAATCTCTTGTCCTCTTTGATTAAAGAAGATTCTCCTTCCAGGGC 317

Qy       1702 ---TAAGCCTGGGATTTGGGCC-----AGAGATAAGAATCCAAACTATGAGGCTAGTTC- 1752
                 |  | ||||||||||||     ||| | |  |  ||||||||||||||||||||
Db        316 CTAAAGCCTGGGGATTTGGGCCCAAGAAGATAAAGAATCCCAAACTATGAGGCTAGTTCT 257

Qy       1753 TTGTCTAACTCAAGACTGTTCTGGAA--TGAGGGTCCAGGCCTGTCAACCAT-GGGGCTT 1809
              |||||||||||||||||||||| |||||   | ||  ||||||||||||||| |||||||
Db        256 TTGTCTAACTCAAGACTGTTTTGGAAATGGGGGTCCCAGGCCTGTCAACCATGGGGGCTT 197

Qy       1810 CTGACCTGAGCACCAAGGTTGAGGGAC--AGGATTAGGCAGGGTCTGTCCTGTGGCCACC 1867
              |||||||||||||||||||||||||||   |  |  | ||||||||||||||||||||||
Db        196 CTGACCTGAGCACCAAGGTTGAGGGACCAGGATTAGGCCAGGGTCTGTCCTGTGGCCACC 137

Qy       1868 TGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATA 1927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 TGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCATA 77

Qy       1928 CTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTAT 1987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 CTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTAT 17

Qy       1988 GAAATAAAAAAAAAAA 2003
              ||||||||||||||||
Db         16 GAAATAAAAAAAAAAA 1


RESULT 12
ABK55110
ID   ABK55110 standard; cDNA; 541 BP.
XX
AC   ABK55110;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human colon cancer-associated cDNA, SEQ ID No 580.
XX
KW   Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
```

```
PN   WO200212280-A2.
XX
PD   14-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US23826.
XX
PR   03-AUG-2000; 2000US-223265P.
PR   02-OCT-2000; 2000US-237406P.
PR   20-MAR-2001; 2001US-277495P.
PR   03-JUL-2001; 2001US-302702P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Pyle RA,  Xu J,  Secrist H;
XX
DR   WPI; 2002-257462/30.
XX
PT   Novel polynucleotide encoding colon tumour polypeptides, useful as
PT   vaccines for treating colon cancers  -
XX
PS   Claim 1; Page 315; 425pp; English.
XX
CC   The invention relates to isolated polynucleotides (I) encoding colon
CC   tumour polypeptides (II). (I) is useful for stimulating an immune
CC   response in a patient and treating colon cancer in a patient.
CC   Oligonucleotides derived from (I) are useful for determining the presence
CC   of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC   compositions, e.g. vaccines, and other compositions for the diagnosis
CC   and treatment of colon cancer. A composition comprising a first component
CC   selected from physiologically acceptable carriers and immunostimulants,
CC   and an antigen-presenting cell expressing (II) is useful for inhibiting
CC   development of cancer in a patient. (I) is useful in the design and
CC   preparation of ribozyme molecules for inhibiting expression of tumour
CC   polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
CC   sequences of the invention.
XX
SQ   Sequence 541 BP; 126 A; 146 C; 141 G; 128 T; 0 other;

  Query Match             25.9%;  Score 533;  DB 24;  Length 541;
  Best Local Similarity   99.1%;  Pred. No. 2.6e-85;
  Matches  536;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1069 GGGCTGGACGAAGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGAC 1128
             || |  | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCGCGTCGACGCCAGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGAC 60

Qy       1129 CACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTT 1188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACCCTACTACACGCTCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTT 120

Qy       1189 CTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGC 1248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCTGCCTCTGAACTGAGCCAAACCTTCAGTGCCTTCCAGAAGGGAGAAAGGCAGAAGC 180

Qy       1249 CTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTG 1308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGTGTGGAGTGTGCTGTGTACACATCTGCTTTGTTCCACACACATGCAGTTCCTGCTTG 240

Qy       1309 GGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGA 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGCTTATCAGGTGCCAAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAGCAAAGGA 300

Qy       1369 GACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTC 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACAATATTCCCTGCTCACAGAGATGACAAACTGGCATCCTTGAGCTGACAACACTTTTC 360

Qy       1429 CATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGA 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACCATAGGTCACTGTCTACACTGGGTACACTTTGTACCAGTGTCGGCCTCCACTGA 420

Qy       1489 TGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCT 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCTGGTGCTCAGGCACCTCTGTCCAAGGACAATCCCTTTCACAAACAAACCAGCTGCCT 480

Qy       1549 TTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGTATCTTGTACCTTTTCAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTC 540

Qy       1609 T 1609
              |
Db        541 T 541


RESULT 13
AAQ89817/c
ID   AAQ89817 standard; DNA; 2562 BP.
XX
AC   AAQ89817;
XX
DT   25-MAR-2003  (updated)
DT   27-NOV-1995  (first entry)
XX
DE   DNA encoding the AUF1 polypeptide.
XX
KW   AU-rich element RNA-binding factor-1; proto-oncogene; c-myc; c-myb;
KW   c-fos; cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             246..1109
FT                   /*tag=  a
XX
PN   WO9511695-A1.
XX
PD   04-MAY-1995.
XX
PF   25-OCT-1994;   94WO-US12273.
XX
PR   29-OCT-1993;   93US-0146421.
XX
PA   (UYWA-) UNIV WAKE FOREST.
XX
PI   Brewer G;
XX
DR   WPI; 1995-178647/23.
DR   P-PSDB; AAR74201.
XX
PT   AUF1 polypeptide and DNA encoding it - limits the expression of a
PT   proto-oncogene that, when over expressed, leads to cancer.
XX
PS   Claim 3; Fig 8; 87pp; English.
XX
CC   The sequence is that of the DNA encoding AUF1 (AU-rich element RNA-
CC   binding factor-1).  AUF1 can be used to treat humans with low levels
CC   of the AUF1 gene, so limiting the expression of a proto-oncogene (
CC   pref. c-myc, c-myb or c-fos) which when over expressed, leads to
CC   cancer.
CC   See also AAQ89818-20.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 2562 BP; 663 A; 608 C; 788 G; 503 T; 0 other;

  Query Match             24.5%;  Score 505.2;  DB 16;  Length 2562;
  Best Local Similarity   94.6%;  Pred. No. 2.7e-80;
  Matches  566;  Conservative    0;  Mismatches   28;  Indels    4;  Gaps    4;

Qy        604 TGTCGCTTTGTCTTCGCTGACCGTGAGAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 663
              || | |||    ||| |      ||   ||||||||||||||||||||||||||||||||
Db       1637 TGACCCTTCTGTTTCTCCCCATGTCCCAGGAAGAAGCTGGTGCTGGAGAACCTGGAGGAC 1578

Qy        664 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGTGGGACAAGCACGCGTGCCCAGCCTAC 723
              |||||||||||||||||||||||||||||||| ||||||||||||  |||||||||
Db       1577 TCCTGCGTGCTGACTGGGCCAGATGATTCCCTGT-GGACAAGCACGCGTCGCCAGCCTAC 1519

Qy        724 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 783
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1518 GTGGGACCTGAGATACTCAGCTCACGGGCCTCATACTCGGGCAAGGCAGCCGATGTCTGG 1459

Qy        784 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 843
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1458 AGCCTGGGCGTGGCGCTCTTCACCATGCTGGCCGGCCACTACCCCTTCCAGGACTCGGAG 1399

Qy        844 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTGCCTGCAGGCCTCTCG 903
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1398 CCTGTCCTGCTCTTCGGCAAGATCCGCCGCGGGGGCCTACGCCTTGCCTGCAGCGCTCTCG 1339

Qy        904 GCCCCTGCCCGCTGTCTGGTTCGCTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 963
                || |  | ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1338 CCCTCGCCGCGCTGTCTGGTTCCGTGCCTCCTTCGTCGGGAGCCAGCTGAACGGCTCACA 1279

Qy        964 GCCACAGGCATCCTCCTGCACCCCTGGCTGCGACAGGACCCGATGCCCTTAGCCCCAACC 1023
                | |||||  |||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1278 G-CACAGCGATCCTCCTGCACCCCTGGCTGCGACA-GACCCGATGCCCTTAGCCCCAACC 1221

Qy       1024 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAAGCC 1083
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1220 CGATCCCATCTCTGGGAGGCTGCCCAGGTGGTCCCTGATGGACTGGGGCTGGACGAA-CC 1162

Qy       1084 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1143
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1161 AGGGAAGAGGAGGGAGACAGAGAAGTGGTTCTGTATGGCTAGGACCACCCTACTACACGC 1102

Qy       1144 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGA 1201
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 TCAGCTGCCAACAGTGGATTGAGTTTGGGGGTAGCTCCAAGCCTTCTCCTGCCTCTGA 1044
```

RESULT 14
ABL87304/c
ID ABL87304 standard; cDNA; 426 BP.
XX
AC ABL87304;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:10282.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US17756.
XX
PR 26-MAY-2000; 2000US-207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide -
XX
PS Claim 1; SEQ ID 10282; 489pp; English.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour cDNA library using well known
CC techniques.
XX
SQ Sequence 426 BP; 105 A; 112 C; 107 G; 102 T; 0 other;

```
  Query Match             20.7%;  Score 426;  DB 24;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-66;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1567 CAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGAC 1626
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CAGAGAAAGGGAGGTATCCCTGTGCCAAAGGCTCCAGGCCTCTCCCCTGCAACTCAGGAC 367

Qy       1627 CCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGA 1686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CCAAGCCCAGCTCACTCTGGGAACTGTGTTCCCAGCATCTCTGTCCTCTTGATTAAGAGA 307

Qy       1687 TTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGC 1746
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TTCTCCTTCCAGGCCTAAGCCTGGGATTTGGGCCAGAGATAAGAATCCAAACTATGAGGC 247

Qy       1747 TAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGG 1806
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TAGTTCTTGTCTAACTCAAGACTGTTCTGGAATGAGGGTCCAGGCCTGTCAACCATGGGG 187

Qy       1807 CTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCAC 1866
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCTGACCTGAGCACCAAGGTTGAGGGACAGGATTAGGCAGGGTCTGTCCTGTGGCCAC 127

Qy       1867 CTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCAT 1926
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CTGGAAAGTCCCAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACAATCCCAGGTCCAT 67

Qy       1927 ACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTA 1986
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 ACTCTAGGTTTTGGATACCATGAGTATGTATGTTTACCTGTGCCTAATAAAGGAGAATTA 7

Qy       1987 TGAAAT 1992
                ||||||
Db           6 TGAAAT 1
```

RESULT 15
AAS37461
ID AAS37461 standard; cDNA; 396 BP.
XX
AC AAS37461;
XX
DT 17-DEC-2001 (first entry)
XX
DE Novel human diagnostic and therapeutic gene #519.
XX
KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
XX
OS Homo sapiens.
XX
PN WO200166753-A2.
XX
PD 13-SEP-2001.
XX

PF 09-MAR-2001; 2001WO-US07787.
XX
PR 09-MAR-2000; 2000US-0188609.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX
DR WPI; 2001-530177/58.
XX
PT New polynucleotides and polypeptides, useful for diagnosis and
PT treatment of breast, lung and colon cancer -
XX
PS Claim 1; Page 727; 1193pp; English.
XX
CC The invention relates to new polynucleotides and polypeptides, useful for
CC diagnosis and treatment of breast, lung and colon cancer. The sequences
CC can be used in detecting differentially expressed genes correlated with a
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample derived from a
CC cell suspected of being cancerous. They can also be used to inhibit
CC tumour growth by modulating expression of a gene product. AAS36943-
CC AAS39338 represent novel human diagnostic and therapeutic coding
CC sequences of the invention.
XX
SQ Sequence 396 BP; 64 A; 138 C; 124 G; 70 T; 0 other;

Query Match 16.0%; Score 330.2; DB 22; Length 396;
Best Local Similarity 97.7%; Pred. No. 1.9e-49;
Matches 335; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          1 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 GCTCTGAGCCCCGGCGGCGCCCGGGCCCACGCGGAACGACGGGGCGAGATGCGAGCCACC 112

Qy         61 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGTTGGAGTTGGATGACAAC 120
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        113 CCTCTGGCTGCTCCTGCGGGTTCCCTGTCCAGGAAGAAGCGGATGGAGTTGGATGACAAC 172

Qy        121 TTAGATACCGAGCGTCCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 180
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        173 TTAGATACCGAGCGTTCCGTCCAGAAACGAGCTCGAAGTGGGCCCCAGCCCAGACTGCCC 232

Qy        181 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 CCCTGCCTGTTGCCCCTGAGCCCACCTACTGCTCCAGATCGTGCAACTGCTGTGGCCACT 292

Qy        241 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGGCCTACCGG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |
Db        293 GCCTCCCGTCTTGGGCCCTATGTCCTCCTGGAGCCCGAGGAGGGCGGGCGGTCCTACCAG 352

Qy        301 GCCCTGCACTGCCCTACAGGCACTGAGTATACCTGCAAGGTGT 343
              |||||||||||||||||||||||    ||||||||||||||||
Db        353 GCCCTGCACTGCCCTACAGGCACTTGAGATACCTGCAAGGTGT 395
```

Search completed: January 16, 2004, 11:32:55
Job time : 595.591 secs